

MODELLING THE DYNAMIC VENTILATORY RESPONSE TO HYPOXIA IN NORMAL SUBJECTS

T.P. Kirby

PhD

University of Edinburgh

1990



CONTENTS

DECLARATION	1
ACKNOWLEDGEMENTS	2
ABSTRACT	4
PREFACE	5
ABBREVIATIONS	7
CHAPTER 1	8
GENERAL INTRODUCTION	8
1.1 AN OVERVIEW OF THE CONTROL OF BREATHING	8
1.1.1 Respiration	
1.1.2 Control of Breathing	
1.2 THE RESPONSE TO HYPOXIA	11
1.2.1 The Carotid Bodies	
1.2.2 The Hypoxic Stimulus	
1.2.3 Other Stimuli to the Carotid Bodies	
1.2.4 Further factors affecting hypoxic response	
1.2.5 Central Hypoxic Depression of Ventilation	
1.3 THE MEASUREMENT OF HYPOXIC SENSITIVITY IN MAN	14
1.3.1 The Response	
1.3.2 Measures of hypoxic drive	
1.3.3 Measurements during transient stimuli	
1.4 FOURIER ANALYSIS	17
1.4.1 The Fourier transform	
1.4.2 Convolution	
1.4.3 The importance of the convolution function	
1.4.4 Discrete Fourier transforms	
1.5 LINEARITY	21
1.6 THE LINEAR DIFFERENTIAL EQUATION AND TIME CONSTANTS	23
1.7 AN OVERVIEW OF MATHEMATICAL MODELS	25
1.8 THE USE OF MODELS IN RESPIRATORY PHYSIOLOGY	26
CHAPTER 2	30
DATA ACQUISITION	30
2.1 THE EQUIPMENT	30
2.2 COMPUTATIONS	32
2.2.1 Expired Tidal Volume (V_T)	

2.2.2	Inspiratory and Expiratory Times (T_I & T_E)	
2.2.3	Breathing frequency (f_B)	
2.2.4	Minute Ventilation (\dot{V}_E)	
2.2.5	Inspired Fractions of CO_2 & O_2 (F_{I,CO_2} & F_{I,O_2})	
2.2.6	End Tidal CO_2 and O_2 (P_{ET,CO_2} & P_{ET,O_2})	
2.2.7	Heart Rate	
2.2.8	Arterial Oxygen Saturation (S_{a,O_2})	
2.2.9	Measurement of Oxygen uptake (\dot{V}_{O_2}) and CO_2 output (\dot{V}_{CO_2})	35
2.3	CALIBRATION	35
2.3.1	Pneumotachograph	
2.3.2	The mass spectrometer	
2.3.3	The mass spectrometer delay	
2.4	SPECIFICATIONS	37
CHAPTER 3		39
IMPROVEMENTS TO THE DATA ACQUISITION PROGRAM		39
3.1	END OF BREATH DETECTION	39
3.2	THE DIGITAL FILTER	42
3.2.1	Choice of filter	
3.2.2	Implementation	
3.2.3	Checking the implementation and finding the delay	
3.3	DETECTION OF SWALLOWING	45
3.3.1	Identifying swallowing	
3.3.2	The most successful algorithm	
3.3.3	The success rate	49
CHAPTER 4		54
THE EXPERIMENTS		54
4.1	INTRODUCTION	54
4.1.1	Testing a system's response from an engineer's point of view	
4.1.2	Physiological and practical considerations	
4.2	METHODS	59
4.2.1	General	
4.2.2	Protocol	
4.2.3	Transient Hypoxia	
4.2.4	Step Hypoxia	
4.2.5	Control of P_{a,CO_2}	
4.3	RESULTS	61
4.4	DISCUSSION	71
CHAPTER 5		72
FOURIER ANALYSIS OF THE DATA		72
5.1	INTRODUCTION	72
5.2	APPLICATION OF FOURIER ANALYSIS TO THE \dot{V}_E - S_{a,O_2} SYSTEM	72

5.3 METHOD	73
5.3.1 Interpolation	
5.3.2 Averaging several waveforms	
5.3.3 The transform	
5.3.4 Finding the transfer function	
5.3.5 Deletion of points	
5.3.6 Plotting the results	
5.3.7 Reconvolution	
5.4 RESULTS	78
5.5 DISCUSSION	82
 CHAPTER 6	 84
INVESTIGATING THE EFFECT OF RATE OF CHANGE OF STIMULUS ON RESPONSE	84
6.1 INTRODUCTION	84
6.2 METHODS	84
6.3 RESULTS	85
6.3.1 A test of resolution	88
6.4 DISCUSSION	89
 CHAPTER 7	 92
A PHYSIOLOGICAL COMPUTER MODEL OF RESPIRATION	92
7.1 INTRODUCTION	92
7.2 THE MODEL	92
7.2.1 Validation of model	
7.2.2 The Ventilation Controller	
7.3 USE OF THE MODEL FOR OUR PURPOSES	98
7.3.1 Adjustments to the model	
7.3.2 Simulating the control of PCO_2	
7.3.3 Studies using the model	
7.4 RESULTS	103
7.5 DISCUSSION	106
 CHAPTER 8	 108
MATHEMATICAL MODELS BASED ON ONE DIFFERENTIAL EQUATION	108
8.1 INTRODUCTION	108
8.2 THE LINEAR DIFFERENTIAL EQUATION	109
8.3 SIMPLE NON-LINEAR TERMS	110
8.3.1 The Saturating Effect	
8.3.2 Threshold effects	
8.3.3 Other non-linear effects	
8.4 METHOD	116
8.5 RESULTS	117
8.6 DISCUSSION	119
8.6.1 Rate of change	
8.6.2 Duration	
8.7 SUMMARY	

CHAPTER 9	124
THE DEVELOPMENT OF A COMPUTER PROGRAM FOR INVESTIGATING A SYSTEM BASED ON TWO DIFFERENTIAL EQUATIONS	124
9.1 INTRODUCTION	124
9.2 TWO DIFFERENTIAL EQUATIONS	124
9.2.1 Inhibition/Potentiation	
9.2.2 Definition of the generalised model	
9.3 OVERVIEW OF THE PROGRAM	128
9.3.1 The aim	
9.3.2 The Program	
9.4 PREPARATION OF THE INPUT FILE	131
9.4.1 Reasons for preparing the file	
9.4.2 The program to prepare input files	
9.4.3 Preparation of S_{a,O_2}	
9.4.4 Preparation of \dot{V}_E	
9.5 THE INPUT TO THE PROGRAM	134
9.5.1 Interpolation of S_{a,O_2}	
9.6 SOLVING THE DIFFERENTIAL EQUATIONS	135
9.6.1 The Fourth Order Runge-Kutta Method for Integration	
9.6.2 Adaptive step size	
9.6.3 Changes required to this routine	
9.6.4 Using the Runge-Kutta routines in the program	
9.7 OBTAINING A VALUE FOR CLOSENESS OF FIT	140
9.7.1 Summary of measure chosen for closeness of fit	
9.7.2 A further use for weighting	
9.8 THE MINIMISATION ROUTINE	144
9.8.1 The Simplex Method	
9.8.2 Adapting the routine for the program	
9.8.3 Problems with minimisation routines	
9.8.4 Problems specific to the simplex method	
9.9 A METHOD FOR SHOWING THAT ONE FIT IS STATISTICALLY BETTER THAN ANOTHER	149
9.10 THE GENERAL APPROACH TO USING THE PROGRAM	150
 CHAPTER 10	 153
THE DEVELOPMENT OF A SUCCESSFUL MATHEMATICAL MODEL	153
10.1 INTRODUCTION	153
10.1.1 Adequacy of fit	
10.2 THE DEVELOPMENT OF A MODEL USING DATA FROM ONE SUBJECT	154
10.2.1 One differential equation	
10.2.2 Two linear differential equations	
10.2.3 Shortcomings of the parallel arrangement	

10.2.4	Cut-off of the output of the slow differential equation	
10.2.5	Addition of the "Saturating Effect"	
10.3	APPLICATION OF THIS MODEL TO OTHER SUBJECTS.....	164
10.3.1	"Inhibition"	
10.3.2	"Potentiation"	
10.4	THE FINAL MODEL.....	169
10.5	PARAMETER ESTIMATION FOR EACH SUBJECT	175
10.5.1	The program	
10.5.2	The method	
10.5.3	Results	
10.6	VALIDATION OF THE MODEL.....	181
10.6.1	Validating the model's fit to each subject	
10.6.2	Conclusion	
CHAPTER 11		186
DISCUSSION		186
11.1	SUMMARY OF THE ARGUMENTS USED IN DEVELOPING THE MODEL.....	186
11.2	THE MODEL IN WORDS.....	189
11.2.1	The two linear differential equations in parallel	
11.2.2	The Saturating Effect	
11.2.3	Inhibition	
11.2.4	Potentiation	
11.3	PHYSIOLOGICAL INTERPRETATION OF THE MODEL.....	193
11.3.1	The fast process	
11.3.2	The slow process	
11.3.3	The saturating effect, inhibition and potentiation	
11.4	THE CONTROL OF ARTERIAL P_{CO_2}	200
11.4.1	Applying these ideas to our results.....	202
11.5	USE OF THE MODEL	203
11.6	IN CONCLUSION - THE FINAL VALIDATION OF THE MODEL.....	204
APPENDIX A		206
THE CONTROL OF P_{CO_2} DURING HYPOXIA		206
A.1	INTRODUCTION.....	206
A.2	BREATH-BY-BREATH \dot{V}_{CO_2}	207
A.2.1	Introduction.....	207
A.2.2	Methods.....	209
A.2.2	Results.....	209
A.2.3	Discussion	211
A.3	A MODEL OF THE P_{A,CO_2} CYCLE DURING BREATHING	212
A.3.1	Introduction.....	212
A.3.2	The Model.....	213

A.3.3 Experiment 1. Factors affecting the PET,CO ₂ -P _a ,CO ₂ difference.....	215
A.3.4 Results	
A.3.5 Experiment 2. Keeping PET,CO ₂ constant using inspired CO ₂	215
A.3.6 Results	
A.4 DISCUSSION.....	221
REFERENCES	224

DECLARATION

The composition of this thesis, and, except where stated, the research described within, are entirely my own work.

T P Kirby

November 1990.

ACKNOWLEDGEMENTS

I would like to express my considerable gratitude to the following people for their help with the research for this thesis and their support during its preparation:

- The late Professor **David Flenley** and the **Wellcome Foundation** who provided the finance and environment for this work to take place.
- The **Staff & Doctors of the Rayne Laboratory**, for their technical help, their time, their encouragement and their conversation.
- Dr **Sue De Cort** and Dr **Jane Hill** who put in many, many hours hard labour obtaining the experimental results quoted in this thesis.
- Dr **Alex Wight** of the Computer Science department, one of my supervisors, for his time and helpful advice (and timely chivvying).
- Professor **Ewart Carson** of the City University in London, who helped us out of several cul-de-sacs, and suggested sensible routes to travel. The model would not have been developed without his help.
- Professor **Ken Saunders** of St George's Hospital in London for inviting me to use his model for a month and for his help once I arrived; and Dr **Chandu Patil** who also provided valuable assistance and ideas.
- Dr **Brian Kirby & Rachel Kirby**, my parents, for their help with the fees, their continuous encouragement and their practical help and criticism during the writing of this thesis.
- **Dorothy May**, who typed up my list of References.
- **BP Exploration**, my current employers, for letting me use their sophisticated equipment, and for letting me have a little study leave when time was tight.

and especially to:

- **Dr Pat Warren**, my physiologist supervisor, who cultivated my meagre knowledge of physiology, walked many treadmill miles while I tested my programs, devoted many hours to reading this thesis, and chivvied me at all the right times.

- **Dr Peter Wraith**, perhaps the only person to understand properly all the work in this thesis, for his many ideas, his computing help, his spotting of the critical mistake, his provision of tools to help in the composition of this thesis, his consistently sound advice and his humour at critical times.

ABSTRACT

The relationship between ventilation and oxygen saturation during hypoxia is usually assumed to be approximately linear, but when the ventilatory response to hypoxia, expressed as ventilation/oxygen saturation, was measured for 10 subjects, in 8 of the subjects it was found to be greater when measured using step change hypoxia than when using transient hypoxia.

Using Fourier techniques, the ventilation/saturation relationship was shown not to be linear. Use of a detailed physiological model failed to reveal a cause for the difference between the response to step and transient hypoxia, and rate of fall of oxygen saturation was shown experimentally and theoretically not to be an important factor in the difference. In the absence of a clear quantitative explanation for the discrepancy, a mathematical model was developed to describe the dynamic ventilatory response.

The model was built by adding terms of increasing complexity to a simple linear differential equation. The simplest model which adequately described the responses of all the subjects consisted of two linear differential equations (1 and 2) in parallel, the input of both being the fall in oxygen saturation, the sum of the outputs giving the rise in ventilation. Equation 1 had a fast time constant (<3 sec), and Equation 2 a slow time constant. Non-linear terms included were a "saturating effect", similar to that described by the Michaelis-Menten formulation, which reduced gain 2 as oxygen saturation fell; and "inhibition" or "potentiation" of gain 1 as the output of Equation 2 increased. This model produced statistically better fits to the data than any of the simpler models tried. As well as providing a more precise description of the hypoxic ventilatory response, the model suggested further experiments that might elucidate the physiological mechanisms occurring during hypoxia.

Using a further model, an appendix discusses flaws in the widely-used technique of attempting to control arterial carbon dioxide tension by using end-tidal carbon dioxide pressure as a controlled variable.

PREFACE

It is fundamental to life in the higher animals that an adequate oxygen supply reaches the muscles that need it, and a complex control system has evolved to match supply to requirements. Modern mankind's lifestyle can, however, put stress on this system. Changing between altitudes alters the amount of oxygen available, and diseases of the lung can seriously reduce the efficiency of oxygen absorption over very long periods. It is important, therefore, to understand as much as possible about the regulatory mechanisms at work.

This thesis describes work carried out in the Rayne Laboratory, University of Edinburgh, between 1985 and 1988, as part of an investigation into one of the mechanisms in the regulation of breathing: the response to hypoxia (ie lower than normal levels of oxygen in the blood) in humans. This was being studied for two reasons. Firstly, there was the simple scientific desire to examine in depth a mechanism which had not been studied elsewhere as intensely as the carbon dioxide regulatory mechanism. Secondly, the group had a medical interest in the subject, for it had been proposed that people with a low response to hypoxia may be more prone to contracting the "blue and bloated" syndrome with emphysema than those with a high response. If this were the case, it may be possible to take preventive action in such individuals before the syndrome took effect, or it may be possible to develop a suitable therapy for those already suffering from it.

The work in this thesis does not directly examine this hypothesis. Instead it attempts to provide an analytical tool by which the hypothesis might subsequently be tested. Such a test requires a reliable and reasonably safe way to quantify the response to hypoxia in members of the general public. Unfortunately, early work in the laboratory showed that two different methods for measuring the response, applied to a single individual, could result in widely different values describing the response. This seemed to conflict with current theory, and prompted a closer examination of these methods and the results. As there was no obvious quantitative physiological theory that could explain the discrepancy, mathematical methods were used to analyse the results, and a mathematical model was

developed that could simulate the results of both methods. Most of this thesis is a description of the steps leading to the development of that model.

The work described in this thesis is far from perfect, and I am very well aware of some of its shortcomings. For instance, the model relies frequently on visual inspection, is based on the data of very few individuals, lacks a satisfactory test of fitted parameter value variance, and the fitting algorithm is slow and far from ideal. These points are discussed in the text, but many of them result from the lack of powerful computing facilities and the lack of time. The thesis itself suffers from having been written after departing from the laboratory: my views on some aspects inevitably altered, and there have been times when I found that I did not have quite the diagram or result necessary for illustrating a point I wished to make.

As with most modern research, this work was carried out as part of a team, and several important ideas originated from discussion with other people. My only part in the collection of the experimental data was in supporting and enhancing the programs and occasionally contributing to discussions of what data was required. Data acquisition has, however, been described in detail in this thesis as it is so fundamental to what follows. The use of Fourier analysis was originally suggested by Dr Peter Wraith, and the basic idea of using differential linear equations with non-linear effects was put forward by Professor Ewart Carson. The details of their use, the solution of many of the problems arising, and most of the associated ideas are, however, my own.

ABBREVIATIONS

CO_2	Carbon Dioxide
$\text{C}_{\text{a},\text{CO}_2}$	Arterial concentration of CO_2
$\text{C}_{\text{v},\text{CO}_2}$	Venous concentration of CO_2
$\text{F}_{\text{I},\text{CO}_2}$	Inspired fraction of CO_2
$\text{P}_{\text{a},\text{CO}_2}$	Arterial CO_2 tension
$\text{P}_{\text{A},\text{CO}_2}$	Alveolar partial pressure of CO_2
$\text{P}_{\text{ET},\text{CO}_2}$	End tidal partial pressure of CO_2
$\text{P}_{\text{v},\text{CO}_2}$	Venous CO_2 tension
$[\text{ET-a}]$	Difference between $\text{P}_{\text{ET},\text{CO}_2}$ and mean $\text{P}_{\text{a},\text{CO}_2}$
f_{B}	Breathing frequency
frc	Functional residual capacity
$\text{S}_{\text{a},\text{O}_2}$	Arterial oxygen saturation
T_{I}	Inspiration time
T_{E}	Expiration time
T_{TOT}	Total breath period
$\dot{\text{V}}_{\text{E}}$	Ventilation
V_{T}	Tidal volume
$\dot{\text{V}}_{\text{CO}_2}$	CO_2 elimination
$\dot{\text{V}}_{\text{O}_2}$	Oxygen uptake
STPD	Standard Temperature and Pressure (Dry)
BTPS	Body Temperature and Pressure (Saturated)

CHAPTER 1

GENERAL INTRODUCTION

This thesis combines work in two broad areas: the physiology of the hypoxic ventilatory response, and the use of mathematical methods and modelling to increase understanding of a physiological system. In this chapter basic overviews of control of breathing and of modelling are provided followed by reviews of work done in these areas. Several of the later chapters use the ideas of Fourier analysis, linearity and linear differential equations, and so a section has been devoted to a basic explanation of these.

1.1 AN OVERVIEW OF THE CONTROL OF BREATHING

1.1.1 Respiration:

Respiration is the process in which oxygen is transported from the atmosphere to the parts of the body where it is needed for oxidation reactions, and waste carbon dioxide (CO_2), created by metabolic activity, is expelled.

In man and other higher animals, breathing brings air into the lungs, where it mixes with any residual air and is brought into close contact with the bloodstream. Oxygen diffuses into the blood, and CO_2 diffuses from the blood into the lungs (and so is expelled during expiration). Oxygen is transported in the blood in two ways: bound to haemoglobin in the red blood cells, and dissolved in the blood plasma. The amount of oxygen bound to haemoglobin is often quoted as a percentage of the maximum possible amount, a value known as **oxygen saturation** (S_{a,O_2}). The amount dissolved in the plasma, the **oxygen tension**, is usually expressed as a partial pressure (P_{a,O_2}). S_{a,O_2} and P_{a,O_2} are related on a one-to-one basis by a sigmoidal (ie S-shaped) curve, which means that during hypoxia S_{a,O_2} does not begin to fall noticeably from its normal level (of about 96%) until P_{a,O_2} is very much reduced. The relationship alters with carbon dioxide tension (P_{a,CO_2}): increased P_{a,CO_2} decreases the affinity of haemoglobin for oxygen and so results in a lower S_{a,O_2} for a given P_{a,O_2} (the Bohr effect).

1.1.2 Control of Breathing:

Changes in metabolic rate, eg due to exercise or digestion of a meal, result in a change in the amount of oxygen required and CO_2 to be eliminated. This is accomplished by changing the pattern of breathing. Increasing **ventilation** (\dot{V}_E , the volume of air leaving the lungs in a minute) increases the volume of fresh air brought into contact with the blood, and so enables an increase in the absorption of oxygen and the elimination of CO_2 . Ventilation is increased by a combination of increasing the **tidal volume** (V_T) of each breath (ie the total volume of air inspired in a single breath) and increasing the **breathing frequency** (f_B , the number of breaths taken per minute). Increasing the heart rate or the stroke volume of each heart beat also increases the amount of oxygen and CO_2 transported.

The level of ventilation is matched to requirements mainly by keeping constant the levels of arterial hydrogen ion concentration (pH) and the partial pressures of CO_2 and oxygen (P_{a,CO_2} and P_{a,O_2}). In healthy individuals at low altitudes oxygen intake and CO_2 elimination remain closely linked, so that maintaining the level of P_{a,CO_2} will tend to maintain P_{a,O_2} automatically. The main ventilatory control uses what is known as a **negative feedback loop** (Fig 1.1). An increase in P_{a,CO_2} or arterial H^+ concentration is detected by receptors on the ventral medullary surface of the brain (the central receptors) and by the carotid bodies in the main arterial system (the peripheral chemoreceptors, described below), and the level of ventilation is increased to cause an eventual reduction in P_{a,CO_2} and H^+ concentration towards their normal levels. In healthy humans at sea level, the detection of **hypoxia** (reduction in level of oxygen) does not play an important part in the control of breathing (Weil et al, 1970), though it becomes important if P_{a,O_2} falls due to a reduction in inspired oxygen (due, for example, to high altitude) or to a mismatch between ventilation and blood perfusion in lung disease. Hypoxia is detected by the peripheral chemoreceptors, and leads to an increase in ventilation in a similar manner to the feedback loop for CO_2 .

The simple negative feedback loops just outlined for maintaining levels of pH, P_{a,CO_2} and P_{a,O_2} are not sufficient in themselves to keep ventilation at the correct level in all circumstances, and some workers (eg Dempsey et al, 1984) have even suggested P_{a,CO_2} may not be a regulated variable. If the

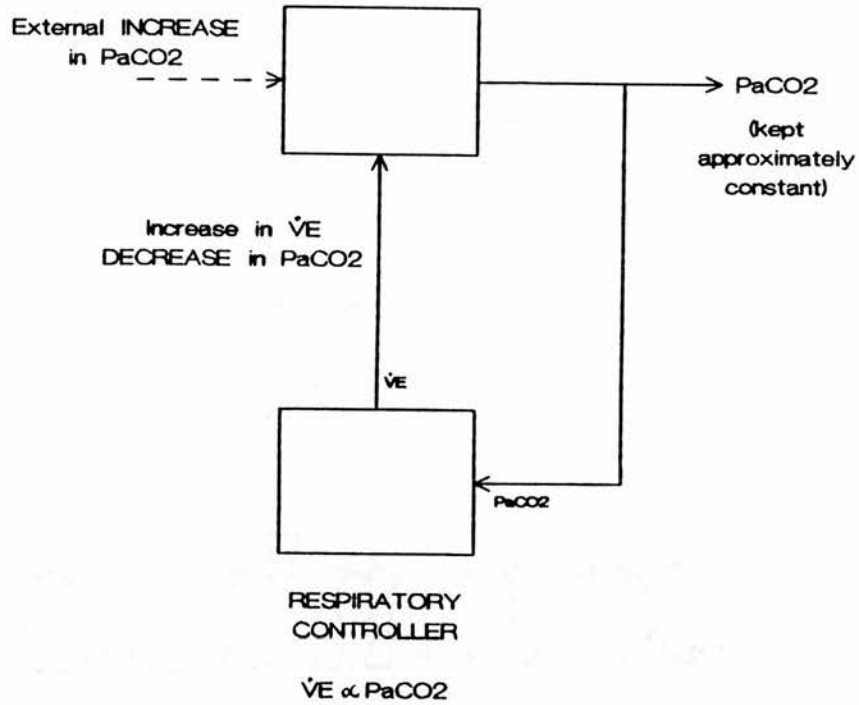


Fig 1.1 The negative feedback loop controlling P_{a,CO_2} . An external change in P_{a,CO_2} is counteracted by a compensatory change in \dot{V}_E . Illustrated is the case when P_{a,CO_2} rises (for example due to increased exercise or inspired CO_2): this is detected and \dot{V}_E rises, thus reducing P_{a,CO_2} towards its normal level.

sole control over ventilation were a feedback loop dependent on P_{a,CO_2} , then during exercise P_{a,CO_2} would be expected to rise to a level above that at rest, but this is not normally found, and indeed it can sometimes fall. The increased ventilation cannot therefore be simply a direct response to raised P_{a,CO_2} . Other stimuli have been suggested, including neural signals from the muscles (experiments have shown that \dot{V}_E is reduced in exercising animals by blocking off nerves returning from hind limb muscles (Kao, 1963; McCloskey & Mitchell, 1972)); excitation of nerves by potassium ions released by the muscles (the increase in \dot{V}_E on starting exercise closely resembles the increase in K^+ ions in venous blood (Tibes, 1977)); neural messages from the brain areas controlling limb movement to those controlling respiration (Krogh & Lindhard, 1913), or coupling between the nerves for limb and thoracic activity (Eldridge et al, 1985); or a link between the changes in cardiac output and the increase in \dot{V}_E (Wasserman et al, 1974). A possible chemical mechanism (suggested initially by Yamamoto, 1960) is that \dot{V}_E increases as the slopes of the breath-by-breath oscillations in P_{a,CO_2} become steeper on exercise, through a feed-forward mechanism (Cross & Semple, 1987). None of these suggested mechanisms provides the whole answer, and several have been disputed by other workers.

1.2 THE RESPONSE TO HYPOXIA

In man, the most important receptors involved in the hypoxic response are the carotid bodies (the peripheral chemoreceptors): this is demonstrated by the disappearance of the hypoxic response in subjects with the carotid bodies removed or denervated (either surgically or chemically) (Guz et al, 1966; Holton & Wood, 1965; Wade et al, 1970; Lugliani et al, 1971; Whipp & Wasserman, 1980). In animals, the aortic chemoreceptors in the ascending aorta and the "miniglomera" in the carotid arteries may also mediate the hypoxic response (Comroe, 1939; Matsuura, 1973), and these may play a part in man in the presence of hypercapnia (Swanson et al, 1978; Honda et al, 1979).

1.2.1 The Carotid Bodies:

These are small ellipsoid structures situated at the bifurcation of the common carotid arteries (two of the main vessels carrying blood to the brain, one on each side of the neck). They are each about 1 cubic millimetre in volume, and have a complex structure made up of two types of cell. They receive an extremely high blood flow of about 2 litres per minute per 100g of tissue (Daly et al, 1954) via the glomic artery, and

this flow is maintained separately from the main blood flow by a sphincter in the artery. The carotid body blood supply is therefore largely unaffected by changes in systemic blood pressure, allowing the carotid bodies to function in adverse conditions.

1.2.2 The Hypoxic Stimulus:

There appears to be a straight-line relationship between \dot{V}_E and S_{a,O_2} (Edelman et al, 1973) and between carotid body activity and S_{a,O_2} (Von Euler et al, 1939). It might be assumed from this that the carotid body responds to oxygen saturation as the main hypoxic stimulus, but this is not the case. Comroe and Schmitt (1938) perfused the carotid body of the dog with blood equilibrated with carbon monoxide (which reduced S_{a,O_2} without reducing P_{a,O_2}) and showed that \dot{V}_E did not increase, so the carotid body must react to falls in oxygen tension rather than in oxygen saturation. This finding has been backed up by numerous studies (eg Duke et al, 1952; Hornbein, 1968; Bartlett & Tenney, 1970; Hebbel et al, 1977). The apparent straight-line relationship between carotid body activity and S_{a,O_2} is due to the response to P_{a,O_2} being roughly hyperbolic, and the dissociation curve relating S_{a,O_2} to P_{a,O_2} being sigmoidal, with the consequence that in the section of interest the convex curve of one approximately cancels out the concave curve of the other to leave a straight line. This relationship is very convenient, and is used throughout this thesis.

1.2.3 Other Stimuli to the Carotid Bodies:

The carotid body responds to increases in P_{a,CO_2} and $[H^+]$ (ie acidity), and it has been shown (Nielsen & Smith, 1952) that the increase in \dot{V}_E due to hypoxia and hypercapnia together is greater than that due to each separately, the same being true for carotid body activity (Eyzaguirre & Lewin, 1961). It is therefore essential to keep P_{a,CO_2} constant during measurements of hypoxic response.

P_{a,CO_2} oscillations: The level of P_{a,CO_2} oscillates with breathing. As fresh air is drawn into the lungs, more CO_2 diffuses out of the blood into the alveoli, and P_{a,CO_2} falls. During expiration alveolar P_{CO_2} rises, limiting further diffusion of CO_2 from the blood, so P_{a,CO_2} increases. The maximum and minimum values of P_{a,CO_2} are determined by tidal

volume, breath duration and the degree of exercise, but the slope of the increase in P_{a,CO_2} during expiration is dependent predominantly on the degree of exercise (Yamamoto, 1960). The carotid body output exhibits fluctuations with the same frequency as the oscillations (Hornbein et al, 1961), but the fluctuations are much larger relative to the mean discharge level than are the fluctuations in P_{a,CO_2} relative to its mean (Biscoe & Purves, 1967; Goodman et al, 1974; Band et al, 1978), which suggests that either there is a threshold in the response to P_{a,CO_2} or that the carotid body is sensitive to the **dynamics of the oscillation** (eg its rate of change (Cross et al, 1984)). If rate of change is important, this has implications for the control of P_{a,CO_2} during hypoxic studies. Adding CO_2 to the inspired gas mixture is likely to affect the slopes of the oscillations, so that even if mean P_{a,CO_2} is kept to its normoxic level, the total stimulus to the carotid body may be altered.

Many **chemicals** circulating in the bloodstream can affect carotid body response. Examples are **catecholamines** (review McDonald, 1981) - noradrenaline can increase hypoxic sensitivity, while dopamine can reduce it; and **potassium ion concentration**, which increases carotid body activity.

Exercise increases the response to hypoxia (Weil et al, 1972), and the carotid body can be shown to play a significant part in the increased \dot{V}_E due to exercise: when subjects breathe hyperoxic gas mixtures (which effectively disable all output from the carotid bodies), \dot{V}_E drops by 15-20% during exercise and 10-15% during rest (Dejours, 1964; Stockley, 1978; Whipp & Wasserman, 1980; Griffiths et al, 1986).

There appears to be an increase in sensitivity to hypoxia in man as **temperature** rises (Petersen & Vejby-Christensen, 1977), and there is evidence in the cat that this might be due to increased sensitivity of the carotid body (Lahiri & Gelfand, 1981).

1.2.4 Further factors affecting hypoxic response:

Hypoxic sensitivity can vary widely between individuals (Dripps & Comroe, 1947; Hirsham et al, 1975; Sahn et al, 1977). Part of this variation may be due to **inherited characteristics** (Collins et al, 1978; Leitch et al, 1975; Kawakami et al, 1982), but other factors found to have an effect on

sensitivity include **state of arousal** (sensitivity is reduced in sleep), **age** (Kronenberg & Drage, 1973; Peterson et al, 1981), **resting metabolic rate** (White et al, 1983), **body size** (Hirshman et al, 1975), and **acclimatisation to altitude** (Forster et al, 1971; Weil et al, 1971).

1.2.5 Central Hypoxic Depression of Ventilation:

If hypoxia is sustained, eventually the ventilatory response diminishes (Holton & Wood, 1965; Weiskopf & Gabel, 1975). Weiskopf & Gabel suggested this might be due to the increased blood flow in the brain resulting from the hypoxia, which causes P_{a,CO_2} to fall and therefore reduces the central drive. Another possibility is that the depression is a manifestation of the biphasic response to hypoxia found in newborn infants (Crosse & Oppe, 1952; Rigatto et al, 1972) which is retained into adult life (Easton et al, 1988) and may be due to central release of a depressant such as adenosine. Whatever the cause, central hypoxic depression seems to depend on the severity and duration of hypoxia, typically coming into effect 5-10 minutes after the onset of hypoxia. It was the desire to avoid this effect which prompted the use of transient stimuli and short duration step hypoxia in the studies described in this thesis.

1.3 THE MEASUREMENT OF HYPOXIC SENSITIVITY IN MAN

One of the main reasons for developing the model in this thesis was to find a reliable and meaningful way of describing and quantifying a subject's hypoxic drive. A variety of methods have been put forward in the past to do this.

1.3.1 The Response:

\dot{V}_E has been used throughout the thesis as if it were the response to hypoxia, but it is not the only variable used as such. The level of activity in the nerves from the carotid body would be the most direct measure of response, and this is frequently measured in animal experiments; but clearly this cannot be used in studies of conscious humans.

One alternative is to use occlusion pressure, the pressure generated by the inspiratory muscles one tenth of a second after the start of inspiration

when the airway is occluded (ie briefly blocked completely) (Whitelaw et al, 1975). Measuring occlusion pressure can, however, disrupt the subject's breathing pattern, and is therefore generally measured every few breaths, which would be unsuitable for transient experiments. It is a measure unaffected by mechanical properties of the lung and chest, and so is more reliable than \dot{V}_E in some diseased subjects whose \dot{V}_E may be distorted by the mechanics of breathing.

Some workers prefer to use mean inspiratory flow, ie tidal volume divided by inspiratory time (V_T/T_I), as a measure of drive. This value describes how large a volume is inspired, and how quickly, these being the two factors that can altered directly under the control of the ventilatory drive. Expiration, at low drives, is largely passive, ie it depends on recoil due to the elasticity of the lungs and mechanical properties of the thorax. \dot{V}_E and V_T/T_I are related simply:

$$\dot{V}_E = V_T / (T_I + T_E) = V_T/T_I \times T_I/(T_I + T_E) .$$

If inspiratory time (T_I) and expiratory time (T_E) change in proportion to each other, V_T/T_I is proportional to \dot{V}_E and the two are effectively equivalent as a measure of response.

The crucial factor in respiration is actually not \dot{V}_E but alveolar ventilation: this must be maintained at the correct level if sufficient CO_2 is to be expelled and oxygen absorbed, and ventilatory drive must ultimately determine that this is so. Measuring alveolar ventilation should therefore be a measure of ventilatory drive, and so long as the **dead space** in the lung (regions not taking part directly in gas exchange) does not change in volume, \dot{V}_E will be proportional to alveolar ventilation. As \dot{V}_E is convenient to measure, it is widely used as a measure of alveolar ventilation, and therefore of ventilatory response, in studies involving subjects with normal lung function. It was the measure adopted in our studies, which were all carried out on subjects with normal lung function.

1.3.2 Measures of hypoxic drive:

Early measurements of hypoxic sensitivity were made using steady state experiments where the \dot{V}_E - P_{a,CO_2} relationships were measured at different levels of P_{O_2} . Hypoxic sensitivity was then expressed as the ratio of \dot{V}_E -

P_{a,CO_2} slopes at two specified levels of P_{O_2} (Nielsen & Smith, 1951; Flenley & Miller, 1967), or as the difference in \dot{V}_E between two curves at a specified level of P_{a,CO_2} (Severinghaus et al, 1966). These measures do not take into account the multiplicative effect of hypercapnia on hypoxia, and the experiments themselves were long and uncomfortable. Lloyd et al (1958) plotted slope (S) of iso-oxic \dot{V}_E - P_{a,CO_2} response lines against P_{O_2} and fitted a hyperbola specified by three parameters (A, D and C)

$$S = D (1 + A/(P_{O_2}-C)) ,$$

adapting this later (Lloyd & Cunningham, 1963) to

$$\dot{V}_E / \dot{V}_E(\text{hyperox}) = 1 + A/(P_{ET,O_2} - C)$$

where $\dot{V}_E(\text{hyperox})$ was the ventilation while breathing a hyperoxic gas mixture, C is a constant (they used 32mmHg) and A is the hypoxic sensitivity. Weil et al (1970) used a similar hyperbola to fit \dot{V}_E - P_{O_2} data from studies in which P_{O_2} was reduced progressively and isocapnically:

$$\dot{V}_E = \dot{V}_E(\text{hyperox}) + A/(P_{ET,O_2} - C) .$$

An alternative to fitting a hyperbola was to fit an exponential curve, eg

$$\dot{V}_E - \dot{V}_E(\text{hyperox}) = \{\dot{V}_E(\infty\text{hypox}) - \dot{V}_E(\text{hyperox})\} \exp(-P_{ET,O_2}/K)$$

where $\dot{V}_E(\infty\text{hypox})$ is the \dot{V}_E for an infinitely large hypoxic stimulus and K is the hypoxic constant (Kronenberg et al 1973).

Use of S_{a,O_2} rather than P_{ET,O_2} as a measure of the stimulus simplifies the mathematics considerably (Cherniack et al, 1971; Rebuck & Campbell, 1974). The formula can be expressed as a simple straight line relationship

$$\dot{V}_E = G \times S_{a,O_2} + \dot{V}_E(0)$$

where G is the hypoxic sensitivity. As discussed in § 1.2.2 above, S_{a,O_2} is not the actual stimulus to the carotid bodies, but the dissociation curve relating S_{a,O_2} to P_{O_2} curves in the opposite direction to that of the curve relating \dot{V}_E to P_{O_2} leaving the \dot{V}_E - S_{a,O_2} relationship approximately a straight line. Using S_{a,O_2} in the formula is therefore just the same as using a hyperbola or fitting an exponential curve: all are just mathematical approximations to the true stimulus-response relationship. S_{a,O_2} ,

however, does have the added advantage of being simple to measure directly at the ear (ie at a point near to the carotid body), continuously and reliably throughout a study. P_{ET,O_2} , used instead of P_{a,O_2} in the earlier formulae, is only an indirect approximation to the stimulus - especially if the alveolar-arterial partial pressure difference is increased due to exercise or disease.

1.3.3 Measurements during transient stimuli:

Progressive isocapnic hypoxia provides many points on a $\dot{V}_E - S_{a,O_2}$ graph, but runs the risk of central hypoxic ventilatory depression (see § 1.2.5) distorting the value of carotid body hypoxic sensitivity measured. Transient stimuli avoid this problem of central depression, but each event provides far fewer points, and so there is a problem of how to analyse these points to arrive at a statistically acceptable result, a problem compounded by the natural breath-by-breath randomness in \dot{V}_E . For example, Edelman et al (1973) plotted the mean of the largest two values of \dot{V}_E resulting from transient hypoxia against S_{a,O_2} ; and Flenley et al (1979) gave 5 hyperoxic breaths during a period of hypoxia and pooled the results of several such tests to improve statistical power in a regression of \dot{V}_E against S_{a,O_2} . The S_{a,O_2} values for each stimulus were shifted relative to the \dot{V}_E values to compensate for the time lag in the measurement of S_{a,O_2} by the oximeter. Work in this laboratory (Gould et al, 1985) has used a method where only the "on" transients in \dot{V}_E and S_{a,O_2} (ie where S_{a,O_2} is falling to its minimum and \dot{V}_E is rising to its maximum) were used to find the $\dot{V}_E - S_{a,O_2}$ relationship. This avoided the distortion in the regression due to the slow return to baseline following the stimulus (probably due to a period of hypocapnia resulting from the increased \dot{V}_E), but resulted in a loss of statistical power and potential problems in interpretation when the number of breaths taken between baseline and maximum \dot{V}_E differed from the the number between baseline and minimum S_{a,O_2} .

1.4 FOURIER ANALYSIS

Fourier techniques are widely used in fields as diverse as image enhancement, filter design, acoustics and astronomy. What follows is a brief description of the concepts used in the discussions of test stimuli

(Chapter 4), digital filter design (Chapter 3), and throughout the whole of Chapter 5.

1.4.1 The Fourier transform:

Any continuous section of waveform (where a measured value is changing with time) is mathematically equivalent to the sum of sine waves of every frequency, each frequency having a particular **amplitude** (height - see Fig 1.2) and **phase** (position in the cycle). Fig 1.3 shows how a simple square wave is built up from an infinite number of sine waves. It is frequently useful to specify a waveform in terms of its component frequencies, and this is achieved mathematically by using a technique known as the **Fourier transform**. This provides a value for the amplitude and phase at each frequency. Fig 1.3e shows the Fourier transform of a square wave where the transform only takes values at discrete frequencies; most Fourier transforms are, however, continuous. The Fourier transform is not just a mathematical device: for example, a prism transforms light waves into their component frequencies, each frequency coming out at a slightly different angle.

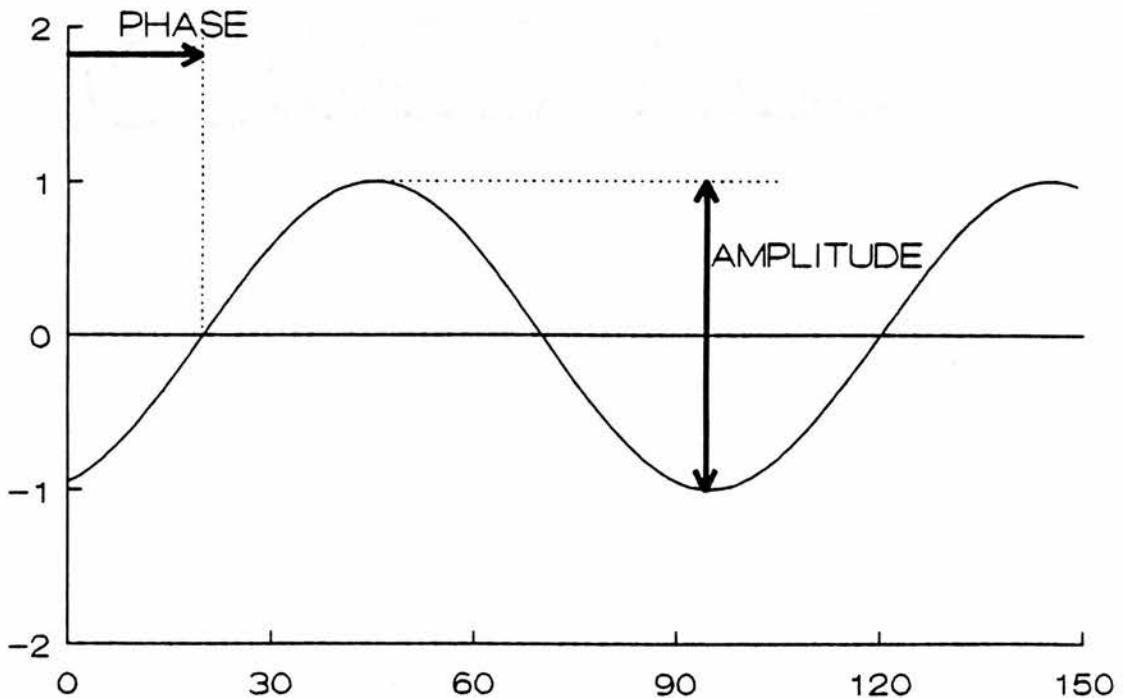


Fig 1.2 Amplitude and Phase of a waveform

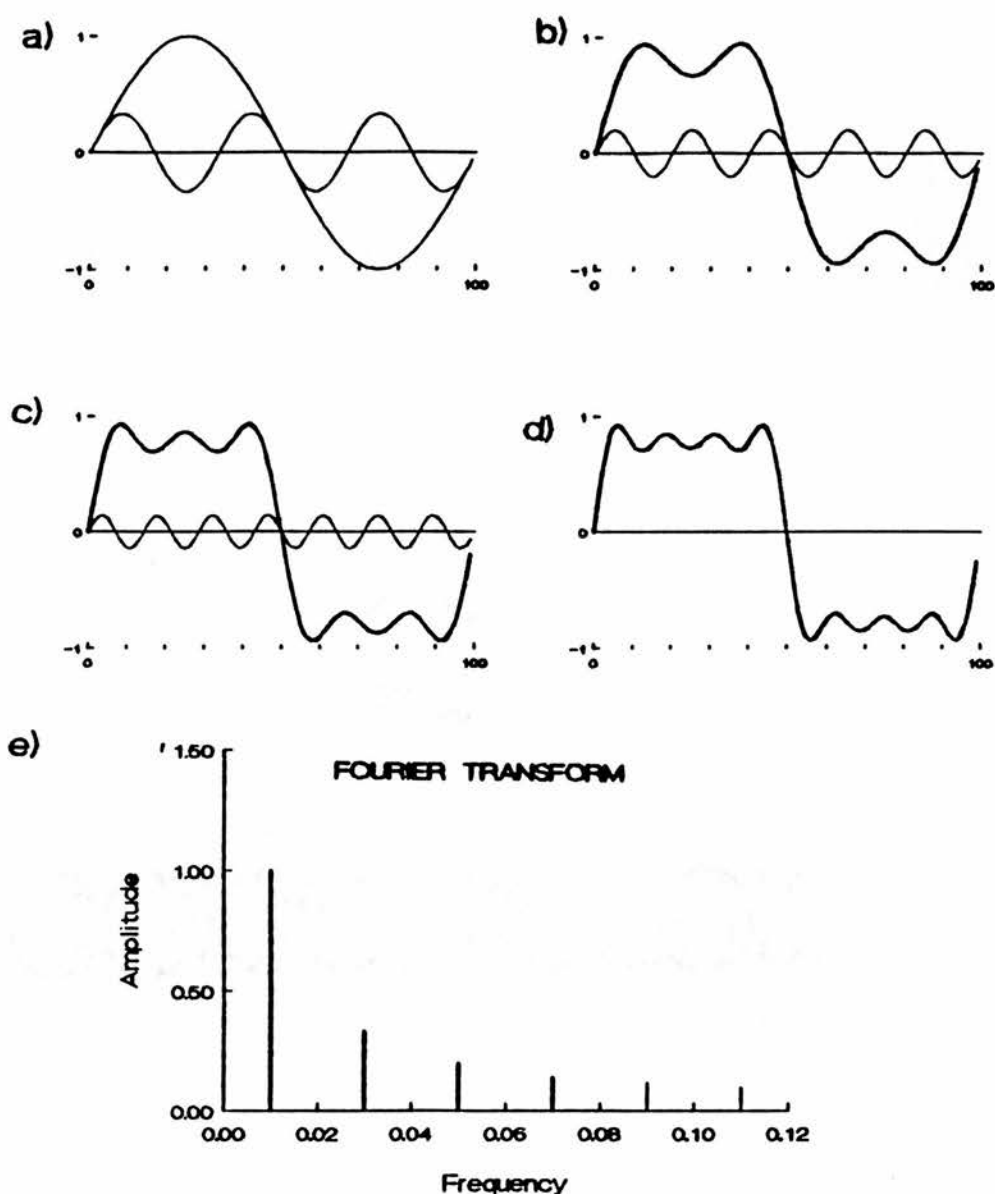


Fig 1.3 Building a square wave from component sine waves. a) to d) show the total of the first four frequencies in the infinite series. The bold curve in b), c) and d) shows the running total of the sine waves shown. e) shows the corresponding Fourier Transform, which, unlike most transforms, only takes discrete values.

1.4.2 Convolution:

The action of filters, lenses, and indeed any linear system which takes an input waveform and distorts it in some way to produce an output waveform, can be described mathematically in terms of convolution. A non-varying function associated with the system, known as the **convolution function**, is combined with the mathematical description of the input waveform via the process of **convolution** to produce the output. Convolution is awkward to deal with when considering an input changing with time, due to the complicated integration involved; but a useful relationship holds between the functions being convolved when working in **frequency space** (ie using the Fourier transforms of the functions involved, working in terms of frequency rather than time). If three functions of time f , g and h are related by convolution (denoted by '*')

$$f * g = h$$

then their transforms F , G and H (all functions of frequency) are related by simple multiplication

$$F \times G = H$$

ie the value of H at a particular frequency is the product of the values of F and G at that frequency. Consider a system with a convolution function g , where f is a waveform entering the system that will result in an output waveform h . The shape of the output can be calculated by transforming f and g to F and G , multiplying these two functions to get H , then performing an inverse Fourier transform to find h in terms of time. A similar process, **deconvolution**, is used to find a system's convolution function if f and h , the input and output, are known. The transforms H and F are divided to obtain G , which is inverse transformed to g .

It is important to stress that the mathematics of convolution holds only for linear systems.

1.4.3 The importance of the convolution function:

The convolution function is essentially a complete mathematical description of the effect a linear system has on an incoming waveform. It is wholly independent of the shape of the input but is a fundamental property of the system itself. It happens also to be the shape of the output if the input is an impulse or delta function (ie an infinitely sharp spike of infinitesimal duration and unit integral), so convolution functions of different systems may be compared as if they were outputs for a standardised input.

The Fourier transform of the convolution function, the **transfer function** or **filter function**, is also a complete description of the effect of the system on an input, but here the description is in terms of how each frequency in the incoming waveform is attenuated and shifted in phase. Electrical filters are usually designed initially by specifying which frequencies are to be reduced by the filter, ie by deciding the shape of the filter function. The digital filter in Chapter 3 was designed in this way.

For a given linear system, the convolution function or filter function can be calculated repeatedly using different shapes of input and will be identical each time (assuming the system itself is not changing between experiments).

1.4.4 Discrete Fourier transforms:

Fourier analysis as described above applies to continuous waveforms, but in many situations the values of input and output are known only at discrete times, for instance in experiments where signals are sampled at regular intervals. A range of techniques, exactly analogous to those for continuous waveforms, has been derived to deal with discretely sampled waveforms. The convolution and filter functions obtained from such data will themselves be discrete. Sampling loses all information about frequencies in the input higher than half the sampling frequency.

1.5 LINEARITY

Convolution function analysis can only be carried out on a system that is **linear**, as the mathematics of convolution assumes linearity. Linearity is

also a concept used throughout the chapters concerning the development of the model.

Linearity is a mathematical term that means more than simply a straight line relationship between two variables (eg $Y = mX$). In a linear system, if the input is multiplied by a constant value, the output will be multiplied by that same value. Furthermore, if two different inputs X_1 and X_2 result in outputs Y_1 and Y_2 , then inputting the two inputs together ($X_1 + X_2$) will result in an output equal to $Y_1 + Y_2$.

A more rigorous definition is: consider an operator L that acts on input X resulting in output Y , ie

$$Y = L(X),$$

then L is a linear operator if:

$$L(aX) = aL(X) = aY$$

$$\text{and } L(X_1) + L(X_2) = L(X_1 + X_2)$$

or, combining these two conditions into one more concise definition:

$$L(aX) + L(bX) = (a+b)L(X) = (a+b)Y$$

where a and b are scalar multipliers of X . Examples of linear relationships are

$$Y = dX/dx$$

$$Y = d^2X/dx^2 + 5.dX/dx - 17.X$$

$$dY/dt + Y = gX$$

which can all be shown to conform to the mathematical definition, while

$$Y = 1/X \int dY/dt + Y = gX^2.$$

are non-linear. In general, a relationship between X and Y that involves X , Y or any of their derivatives being raised to a power other than 1 will be non-linear.

1.6 THE LINEAR DIFFERENTIAL EQUATION AND TIME CONSTANTS

It may not be immediately obvious that the final example on the list of linear equations in the last section conforms to the definition of linearity, and as it forms the basis of much of the modelling work described in later chapters it is worth showing this explicitly, using a more general form including a multiplier τ . If

$$\tau \frac{dY}{dt} + Y = gX$$

$$\text{and } \tau \frac{dy}{dt} + y = gx$$

then it is easy to show that the equation still holds after replacing X by aX and Y by aY to fulfil the first condition

$$\tau \frac{d(aY)}{dt} + (aY) = g(aX);$$

and summing the two equations so that the input is $(X+x)$ leads to an equation which still holds when Y is replaced by $(Y+y)$

$$\tau \frac{d(Y+y)}{dt} + (Y+y) = g(X+x).$$

The simple linear differential equation relates two quantities that vary with time, X , which may be thought of as the stimulus, and Y , the response. In steady state $dY/dt = 0$, and this equation becomes

$$Y = gX.$$

In other words, the value of Y (if X is kept constant for sufficient time to reach steady state) will be proportional to the value of X . This simple proportionality will hold approximately when τ is small or dY/dt is small as the left hand term will then be unimportant. When this is the case the response will simply follow the shape of the stimulus but will be magnified by a factor g , often referred to as the **gain** of the relationship (Fig 1.4b).

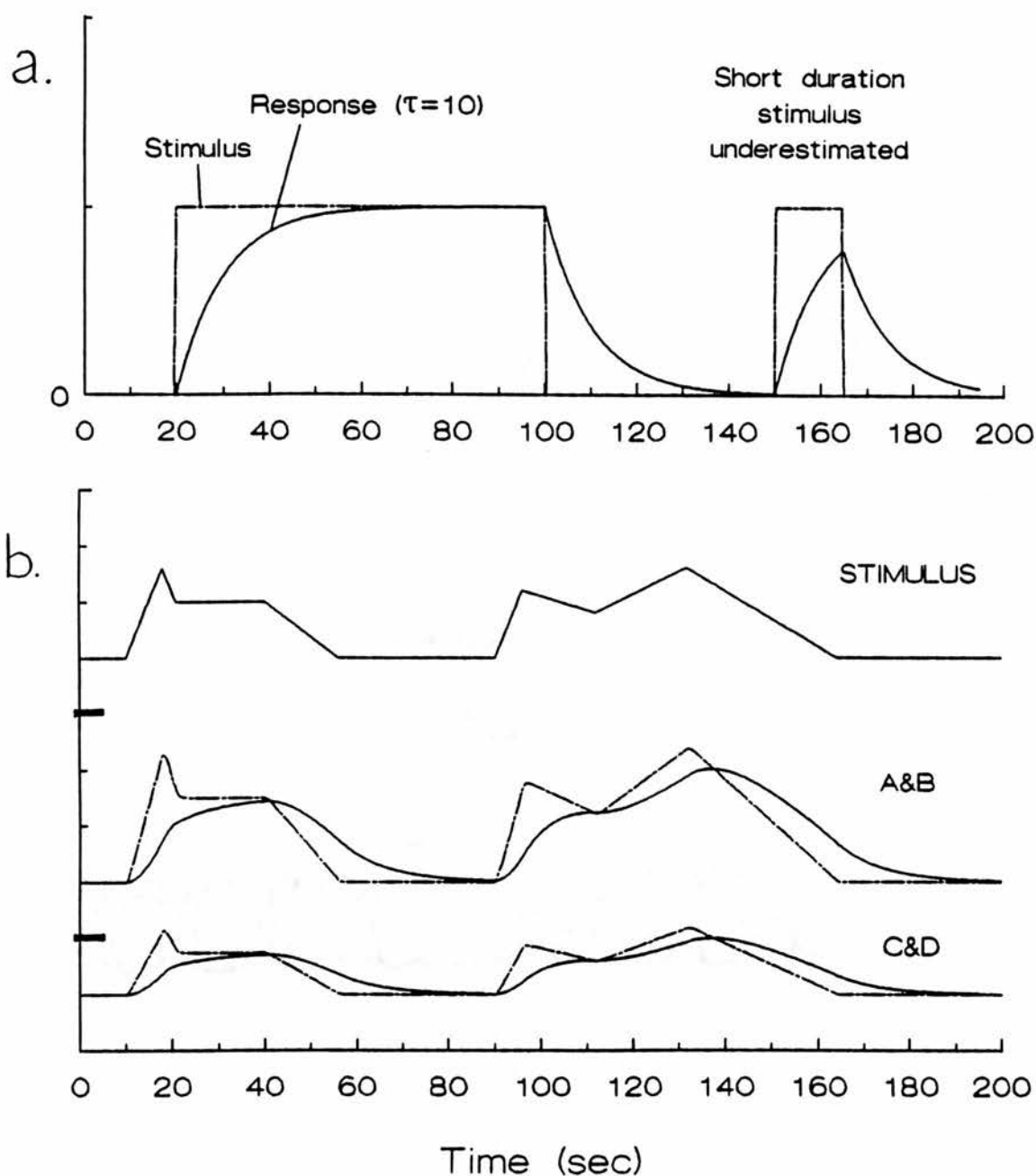


Fig 1.4 The Response of a Linear Differential Equation. a) shows the response with a gain of 1 and a time constant of 10 seconds to a long and a short step stimulus, demonstrating the exponential timecourse and the underestimation of the response to the shorter stimulus. b) shows a the effect of different gains and time constants on the response to a stimulus: curves A and B have gain 1.5, C and D have gain 0.75, the dashed curves have time constant 1 second, and the continuous curves have time constant 10 seconds.

When dY/dt or t are not small, the solution to this type of equation for a given time-course of X involves changes in Y which are dependent not only upon X but also on t , the **time constant**. The larger the time constant, the longer it takes Y to reach its steady state value after a change in X . One consequence of this is that if a change in X is short-lived, the associated change in Y may not have time to reach its full response (Fig 1.4a). The form of the solution has Y changing along an exponential time-course from its initial value just before the change in X to its resting value when steady-state is reached (Fig 1.4a). Fig 1.4b shows examples of the response of equations with different values of t and g to a stylised stimulus.

1.7 AN OVERVIEW OF MATHEMATICAL MODELS

A model, in the sense used in this thesis, is basically just a mathematical formulation which describes systems or events in the real world. Models can be created and used for many reasons, such as:

- Summarising a large number of data points succinctly as an equation, to help examine trends or to predict new values.
- Formulating theories about a set of processes into a series of equations to check that their interactions have been properly understood, and possibly to help test new related theories.
- Modelling a well-understood system mathematically to allow the estimation of parameters that cannot be measured directly.

Models can be extremely complex, but need not be. Perhaps the most frequently used model is also the simplest: in all branches of science straight lines are superimposed on data as a convenient aid to analysis and discussion. The model being used is the simple relationship

$$Y = mX + c,$$

often with the added model that the data points, through measurement or other inaccuracies, lie normally distributed about their "true" values. At the other end of the complexity scale come the large scale meteorological models which require supercomputers to solve the mass of equations encoded in them.

There are certain things in common to all model construction. A model must have a **purpose**. During formulation of the model the degree of approximation that is appropriate must first be decided, and then mathematical equations which describe the system to that **degree of approximation** should be chosen. The model needs some sort of **validation** - a check that it is fulfilling its purpose in an appropriate way. These concepts are discussed more formally by Murray-Smith & Carson (1988), and model validation, in particular, will be considered in depth in Chapters 10 & 11 in this thesis; but the ideas can be demonstrated in action even with the simple straight line model described above. There is no point in fitting a straight line to data without a clear purpose in mind. Frequently the data will not lie exactly on a straight line due to measurement inaccuracies, and although a more complex model might involve detailed descriptions of the state of the measuring instrument as each point was measured to allow a more precise fit to the data, it is usually felt that this degree of accuracy is unnecessary, or even undesirable, and that the simple straight line model is adequate. Validation is also essential: a best straight line can be drawn even through a completely random set of points, but would clearly be worthless; linear regressions are therefore generally accompanied by statistical measures of their validity.

1.8 THE USE OF MODELS IN RESPIRATORY PHYSIOLOGY

Modelling has been an important tool in understanding respiratory physiology for a long time, and its use is widespread (eg Cramp & Carson (1988); Khoo & Yamashiro (1989); Whipp & Wiberg (1983)). Modelling in this field has been used for a wide variety of purposes, including diagnosis, experimental design and dose/response models (Carson & Murray-Smith (1988)). This section will list just a few of the models of more relevance in the field of the control of breathing.

Perhaps the earliest model in this field is found in the work of Haldane & Priestley (1905), who discovered the existence of the negative feedback control loop regulating P_{A,CO_2} . Gray (1946) proposed that P_{A,CO_2} , P_{O_2} and pH each affected respiration **independently** and **additively**, the "multiple factor theory", and for 20 years this assumption was built into respiratory

models. The first attempt to use modelling to examine the dynamics of the chemical control of respiration was that of Grodins et al (1954). Their model was very simple and consisted of simulations of just two compartments, a lung reservoir through which gas flowed constantly, and a tissue reservoir. The flow of gas through the lung was determined by a feedback loop dependent on the P_{CO_2} in the tissue compartment. This model proved very successful at simulating inspired CO_2 steps, in particular the difference between the rate of increase of ventilation at the start of the step and rate of decrease at the end.

The simple model of Grodins was gradually expanded by various workers, and more physiological details were added over the years. First Defares et al (1960) added a brain compartment where measurement of P_{CO_2} levels took place for the feedback loop. Then Horgan and Lange (1963) added in an independent P_{O_2} control loop and time lags between the lung and brain compartments, later expanding the model by dividing the brain compartment into 3 sections representing cerebrospinal fluid, the brain CO_2 reservoir, and the ventrolateral surface where P_{a,CO_2} was detected by the controller (Horgan and Lange, 1965). The controller also responded to the level of P_{O_2} and P_{CO_2} in the arterial blood, and the various time constants were derived from data collected in cats. In the same year Milhorn et al (1965) published a model in which cerebral blood flow depended on P_{a,CO_2} and P_{a,O_2} , and the controller used arterial P_{O_2} and brain P_{a,CO_2} .

Grodins et al (1967) took the controller to a further level of complexity by having it respond to pH and P_{O_2} at the carotid bodies (ie the values were taken at the transport time from lung to carotid bodies taking into consideration changes in blood flow), and pH in the cerebrospinal fluid, again with the appropriate time lag. This resulted in a two component response, but unfortunately there was a sharp break as the second component began to take effect, and the response to step changes of hypoxia displayed apnoea. Milhorn et al (1972) improved the two component response to pH by having their controller respond to the pH at a site between the cerebrospinal fluid and the deep brain, assuming that CO_2 was diffusing between the two along an exponential P_{CO_2} function. An oxygen control loop was added a year later (Milhorn and Reynolds, 1973), but was not compared directly with actual data.

Lloyd et al (1958) proposed a multiplicative interaction between the response to P_{O_2} and P_{CO_2} , but this did not find its way into models (apart from Longobardo et al's (1966) model of Cheyne-Stokes breathing) until Duffin (1972) introduced (into an analogue computer model) a more complex interaction between the response to P_{CO_2} and P_{O_2} at the carotid body, which, although it was based on pharmacological reasoning, was similar to the multiplicative interaction described by Lloyd et al. Yamamoto and Hori (1971) were the first to concentrate properly on the effects of breathing (rather than assuming the lung was simply a vessel through which there was a continuous flow of gas), and they included terms for chest mechanics which allowed the model to convert respiratory drive into an instantaneous airflow. They did not, however, have an oxygen controller in the model. The multiplicative interaction between the responses to P_{a,CO_2} and P_{O_2} and the respiratory cycle play important parts in the model of Saunders et al (1980), and a fuller description of this model will be given in Chapter 7. A more complex "breathing" model was developed by Longobardo et al (1982) to examine hypotheses regarding apnoea in sleep. Taking breathing into account in a different way, Poon (1983) proposed that the controller minimised a function dependent on the respiratory work required as well as on the chemical stimulus.

The models listed so far have all been of the same sort - direct mathematical representations of various physiological processes occurring in compartments that are supposed to mirror parts of the human body. An alternative approach is to model the response as a whole rather than splitting it up into physiological components. The steady-state relationships between \dot{V}_E and P_{O_2} or S_{a,O_2} described earlier are examples of this. Another form of this, but bringing in dynamic aspects, is the modelling of the system as a single "black box", the workings of which are summarised as a "transfer function" (see §1.4.1). This approach, with examples, will be considered in more depth in Chapter 5. The transfer function summarises all responses of a black box concisely but is awkward to interpret in any physiological way; another approach might be to summarise the dynamics of the black box as a set of equations relating output to input. This is effectively just an extension of the simple steady-state equations already described. An example of this applied to the response to CO_2 is given by the two-compartment model of Swanson and

Bellville (1975), extended by Wiberg et al (1979). In this model, the breath-by-breath ventilation is the sum of a peripheral and a central component, each showing linear dependence on the difference between delayed P_{ET,CO_2} and a baseline level, each component having a separate time constant and gain in a simple linear differential equation. In the extension to the model, Wiberg et al added in a term that caused the time constant of the central component to change with cerebral blood flow, itself linearly dependent on P_{a,CO_2} . Dutton et al (1973) derived a model for the dynamic response to hypoxia in dogs. Like the Swanson and Bellville model, this model involved a fast time constant component and a slow time constant component, but here the differential equations were second-order, including a term in the response dependent on the rate of change of P_{a,O_2} . The equations were derived from data obtained by manipulating directly the P_{O_2} of blood perfusing the carotid arteries of dogs, and so much more rapid rates of change were obtainable than would be the case in human studies where P_{a,O_2} can only be altered indirectly by changing inspired gas mixtures. If a similar rate of change component exists in the human response, it is unlikely that it would be noticeable in such studies.

The great majority of the models described are concerned primarily with the CO_2 feedback loop, with the hypoxia controller often added in more as an afterthought. In this thesis, several of the different types of model are used in an examination specifically of the response to hypoxia. Transfer functions are calculated in Chapters 5 and 6, a multi-compartment model is adapted in Chapter 7, and then in Chapters 8, 9 and 10 a parametric model based round differential equations, strikingly similar to those just described of Swanson and Bellville and of Dutton et al (without the rate of change component), is developed to summarise the dynamic hypoxic response.

CHAPTER 2

DATA ACQUISITION

This chapter describes the experimental set-up and the calculations performed to obtain the basic data used in the studies.

2.1 THE EQUIPMENT

The equipment used in the main studies is summarised in Fig 2.1, and technical specifications (manufacturers, sizes etc) are given in the final section of the chapter.

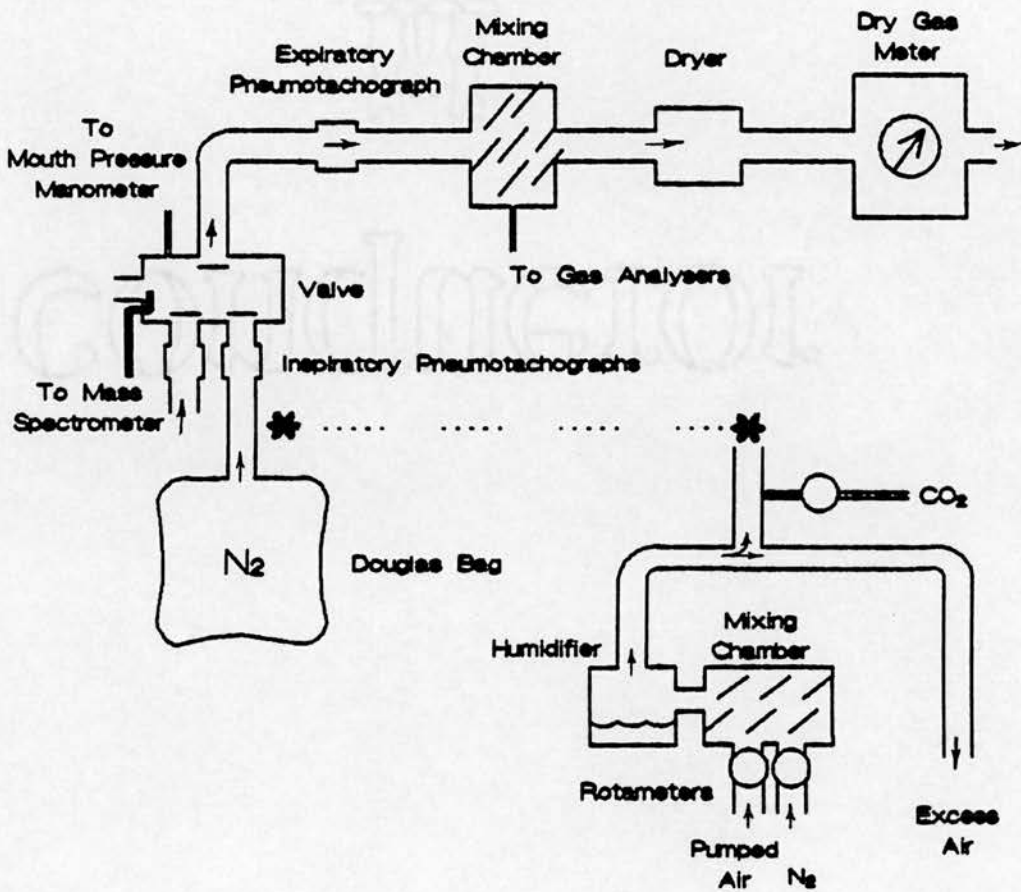


Fig 2.1 A Schematic Representation of the Equipment Used in Data Acquisition. During step hypoxia, the Douglas bag was replaced by the equipment shown to the right.

The subjects walked on a level **treadmill**, and the speed of this was adjusted so that the subject exercised where possible at a set \dot{V}_{O_2} (1 l/min). The subject wore a noseclip and breathed into a **mouthpiece** attached to a custom-built low-resistance **valve** into which inspired gas entered from one of two inspiratory ports and expired gas left via a third.

Pneumotachographs attached to these ports measured inspiratory and expiratory flow; and changes in pressure in the valve (the "mouth pressure") were measured by a **micromanometer** via a narrow tube containing a resistance to dampen down "noisy" pressure fluctuations. A **mass spectrometer** sampled the gas continuously via a heated probe at a point close to the end of the mouthpiece. In later studies, a **five-way valve** was added to the inspiratory port of the custom-built valve to allow rapid switching between a variety of inspired gas mixtures.

The inspired gas: Before using the five-way valve in the circuit, one of the custom-built valve's inspiratory ports allowed the inspiration of room air, whilst the other was attached either to a Douglas bag containing 100% N_2 for the transient hypoxia experiments, or to a rotameter system during step hypoxia experiments. Three **rotameters** were used to control the relative flow rates of nitrogen and CO_2 (both from cylinders), and pumped room air. The flows of air and nitrogen were adjusted to give the required fraction of oxygen in the mixture, and the resultant gas passed at a high flow rate (100 l/min) through a mixing chamber and into a 5cm diameter tube. A T-piece in this tube was attached to the inspiratory port of the valve to allow the gas mixture to be inspired, and the tube continued for 2 metres beyond the T-piece to prevent room air affecting the inspired mixture. This arrangement, with the high flow rate, ensured that the gas supply exceeded respiratory requirements without putting any positive pressure on the inspiratory valve. CO_2 entered the system near to the T-piece: this was used to control end tidal P_{CO_2} on a breath-by-breath basis. When the five-way valve was used, room air, the Douglas bag and the rotameter system T-piece were connected to different ports, and the second inspiratory port of the custom-built valve was not used.

The **expired gas**, after passing through the pneumotachograph, entered a **mixing box** to ensure even mixing and then was cooled to condense out water vapour. The resultant dry gas passed through a **dry gas meter**, which

sent a signal to the computer every 10 litres of gas for calibrating the pneumotachographs (see below). Expired gas was sampled from the mixing chamber at 500ml/min by oxygen and CO₂ analysers for the calculation of steady state gas exchange.

Arterial oxygen saturation was measured using an ear oximeter; and an electrocardiogram (ECG) was used to monitor heart rate.

The phase of respiration was detected by a digital inspiration/expiration switch set by electronic circuitry that reacted directly to the mouth pressure signal. When the mouth pressure was greater than a small threshold value above atmospheric pressure, the switch registered "expiration". When it dropped below a threshold value below atmospheric pressure, the switch changed to "inspiration". The threshold levels were established experimentally to reduce spurious fluctuations between "inspiration" and "expiration" due to noise when the pressure was at approximately atmospheric pressure. During normal breathing the mouth pressure passed the thresholds rapidly at the start of inspiration or expiration, but there was inevitably a small delay; and this, coupled with the resistance in the manometer tube (needed to dampen down noise, but which again caused a loss of precision over the start of a new breath), led to the development of an alternate method of detection, which will be described in Chapter 3.

2.2 COMPUTATIONS

The computer, running a custom-written data acquisition program, sampled the following signals (at 60 Hz - but see §3.2.1):

- inspiratory and expiratory flows (pneumotachographs)
- mouth pressure (manometer)
- O₂ and CO₂ fractions near the mouth (mass spectrometer)
- oxygen saturation (ear oximeter)
- heart rate (ECG)
- temperature of dry gas meter

- inspiration/expiration switch (electronic)
- data storage ON-OFF & event marker (experimenter)

Chosen signals were displayed graphically on a display screen, and other signals plotted continuously on a printer. At the end of each expiration the following values were calculated (and stored on disk if required):

2.2.1 Expired Tidal Volume (V_T):

The sampled voltages from the expiratory pneumotachograph signal were summed during the course of each expiration, and the total converted to a volume by multiplying by a calibration factor (obtained using the dry gas meter, as explained below in §2.3.1) and correcting the result to BTPS (Body Temperature and Pressure, Saturated).

2.2.2 Inspiratory and Expiratory Times (T_I & T_E):

These were the times between changes in the inspiration/expiration switch, ie expiratory time was the period from when the mouth pressure signal crossed the positive threshold to when it dropped below the negative threshold at the start of the next inspiration. This meant that pauses (periods where the subject was neither inspiring nor expiring) were always part of the preceding half breath: a pause was end-expiratory rather than pre-inspiratory.

2.2.3 Breathing frequency (f_B):

This was calculated at the end of each expiration as the reciprocal of the total breath period, multiplied by 60 to get breaths per minute, ie

$$f_B = 60 / (T_I + T_E)$$

2.2.4 Minute Ventilation (\dot{V}_E):

This was calculated breath-by-breath as the product of breathing frequency and tidal volume, giving a result in litres per minute (BTPS). (The "E" denotes the use of expiratory rather than inspiratory tidal volume in the calculation.)

2.2.5 Inspired Fractions of CO₂ & O₂ (F_{I,CO₂} & F_{I,O₂}):

These values were measured in the period 150 to 450 ms after the start of inspiration, or from 150 ms to the start of expiration if inspiration lasted less than 450 ms. During that time the signals from the mass spectrometer were averaged, and the results multiplied by a calibration factor worked out immediately before the study (see §2.3.1) to obtain the fractions as percentages. The signals were sampled only during this period for two reasons: the 150 ms allowed time for any previously expired gas left in the valve to be inhaled; and the inspired fractions changed so little that no more information was gained after 450 ms, and the time was better used by the computer in performing other necessary calculations.

2.2.6 End Tidal CO₂ and O₂ (P_{ET,CO₂} & P_{ET,O₂}):

Rather than trying to find the values of P_{CO₂} and P_{O₂} that occurred immediately before the change from expiration to inspiration, the data acquisition program found the highest value of P_{CO₂} during expiration, and noted the P_{O₂} at that time. This was a more reliable approach, because any delay in (or anticipation of) the detection of the end of expiration would have a marked effect on the measured end tidal values. Partial pressures were calculated by multiplying the mass spectrometer signals by the same calibration factor as for the inspired fractions, to get values as percentages; and then these values were corrected to BTPS.

2.2.7 Heart Rate:

The program measured the time for 8 digital pulses from the ECG (triggered by the subject's R-waves), found the average and divided this into 60 to get an average heart rate in beats per minute (bpm). This resulted in a smoothed, effectively continuous signal. The current value of this at end expiration was stored.

2.2.8 Arterial Oxygen Saturation (S_{a,O₂}):

The colour of blood changes with degree of oxygen saturation, from bright red when fully saturated to deep purple when completely desaturated. The ear oximeter shines a light through the subject's earlobe, using the relative levels of absorbance of several wavelengths to calculate the colour of the

blood, and thus the saturation. This provides a direct, convenient and non-invasive method of finding the S_{a,O_2} at a site near the carotid bodies. Unfortunately, the signal output by the oximeter was not ideal. The output could only take integer values, so when the S_{a,O_2} being measured was halfway between two integers, a rapid fluctuation between the two values occurred in the output. The output was further distorted as the oximeter reacted with a delay and a time constant (measured in 1979 to be 0.34s and 1.61s respectively (Douglas et al, 1979)). The conversion constant from oximeter voltage to S_{a,O_2} was obtained by reading the voltages output for 0% and 100% S_{a,O_2} by the oximeter in its calibration mode, and then using linear interpolation. Values for S_{a,O_2} were stored as the mean over a breath.

2.2.9 Measurement of Oxygen uptake (\dot{V}_{O_2}) and CO_2 output (\dot{V}_{CO_2})

These values were measured over two 2-minute periods starting 5 minutes and 7.15 minutes after the onset of rest or exercise, to establish that steady state gas exchange had been reached (if the results of the two periods differed by more than 100 ml/min then the measurements were repeated over further periods until agreement between successive measurements was achieved). During the two minutes the volume of expired gas passing through the dry gas meter was used to calculate mean ventilation, and three readings from the O_2 and CO_2 analysers (sampling gas from the mixing box) allowed the average gas concentrations, and therefore the rate of gas uptake or output (in l/min).

2.3 CALIBRATION

2.3.1 Pneumotachograph:

The calibration factor for converting the summed signal from the expiratory pneumotachograph to a volume in litres was initially calculated immediately prior to the start of a study. Air was pumped (using an animal respirator which simulated breathing) through the pneumotachograph and the attached dry gas meter, and the output signal was sampled exactly as during the study. The dry gas meter output a digital pulse as every 10 litres of gas passed through it. The pneumotachograph signal was summed between two such pulses and divided into 10 to get the calibration factor, then corrected to BTPS. Although

pneumotachographs measure flow, this calibration effectively converts directly to a volume. It was unnecessary to calculate instantaneous flow, as what was of importance was the expired tidal volume, from which an average flow over a breath (\dot{V}_E) was calculated.

The relationship between the flow through the pneumotachograph and the voltage output could change with time, for example due to temperature changes during a study. To prevent this introducing systematic errors into measurements of V_T and \dot{V}_E , the calibration factor was re-calculated throughout the study every time a pulse from the dry gas meter indicated that another 10 litres of gas had been expired. The calibration factor used was the average of the last 3 calculations to smooth out sudden changes.

The inspiratory pneumotachograph could not be calibrated directly, but if there were no leaks in the valve between the inspiratory and expiratory pneumotachographs (and care was taken that there were none), and it was assumed that subject's residual lung capacity remained constant, then the same dry gas meter pulses could be used with the output from the inspiratory pneumotachograph to obtain a second calibration factor.

2.3.2 The mass spectrometer:

The mass spectrometer output 4 voltages, corresponding to the fractions of nitrogen, oxygen, carbon dioxide and argon in the sampled gas. Both the mass spectrometer and the computer program required the sampling of a selection of gas mixtures to calculate calibration factors for converting the voltages to gas fractions. Six commercially prepared gas mixtures of different (known) concentrations of N_2 , O_2 , CO_2 and A were sampled sequentially, and the mass spectrometer voltages associated with each gas were compared with the corresponding gas fractions, using linear regression to obtain the calibration factors. The calibration was only accepted if the uncertainty in each factor was less than 1%.

In practice the argon value was so small that it was ignored by the program. Once the equations had been solved, the calibration factors were identically scaled so that the total fractions for the four gases summed to 100% (this adjustment was generally small).

2.3.3 The mass spectrometer delay:

The time taken by the mass spectrometer in transporting sampled gas from the probe to the analysers and then in analysing it was measured before each study by switching the probe very rapidly between sampling a gas mixture containing a high proportion of CO₂ and sampling room air. The switch was done using a spring mechanism which simultaneously sent a signal to the computer. The time was measured from this pulse to when the CO₂ fraction (as detected by the mass spectrometer) was halfway between the gas mixture value

This delay was then used by the data acquisition program throughout the study. All signals read by the computer except for those from the mass spectrometer were stored in a buffer for a time equal to the delay before being used, thus ending up exactly synchronised with the mass spectrometer outputs.

2.4 SPECIFICATIONS

- **Custom-built valve:** Dead space: 120 ml (including mouthpiece)
Resistance: 0-4.9 cmH₂O over flow range 0-110 l/min. Operation: Switch from one inlet port to the other using a hand-operated lever-operated spring mechanism.
- **Five-way valve:** Make: Hans-Rudolph 2440 series, Gatlin shape valve.
Dead space: 95 ml Resistance: 0-4.5 cmH₂O over flow range 0-110 l/min.
Operation: Valves closed by pneumatically operated balloons.
- **Pneumotachographs:** Make: Fleisch No. 2
- **Mouth pressure manometer:** Make: Furness Controls FCO/4
- **Ear Oximeter:** Make: Hewlett-Packard 47201A with fibreoptic earprobe
Delay: 0.3 sec Response: time constant 1.6 sec (fast mode)
- **Electrocardiogram:** Make: Hewlett-Packard 78351A
- **Mass Spectrometer:** Make: VG-Medical Spectralab M Sampling: 20-25 ml/min via a heated probe about 7 cm from the subject's mouth Transit time: 100-200 ms Response: 65 ms (10% to 90%)

- **Oxygen Analyser:** Make: Sybron-Taylor Servomex Oxygen Analyser 570A
Sampling: For this and CO₂ Analyser - 500 ml/min
- **CO₂ Analyser:** Make: Gould-Godart Capnograph Mark III
- **Rotameters:** Make: Rotameter Manufacturing Co. Ltd. Flows: Air: 0-200 l/min; N₂: 0-50 l/min; CO₂: 0-10 l/min Mixing Box: 10.5 l
- **Mixing Box (expiratory):** Volume: 3.2 l
- **Dry Gas Meter:** Make: Parkinson-Cowan CD4
- **Computer:** Make: Digital Equipment Corp. PDP 11/73 Language: PDP Macro (assembler) Programs: Custom-written ADC's: All 12-bit

CHAPTER 3

IMPROVEMENTS TO THE DATA ACQUISITION PROGRAM

This chapter deals with improvements and additions made to the data-acquisition computer program in three areas: more precise detection of the start and end of inspiration and expiration, the implementation of an on-line filter, and an on-line method for detecting when swallowing occurs.

3.1 END OF BREATH DETECTION

It is important to locate as precisely as possible the time when expiration ends and inspiration begins, or vice versa. This information is used to calculate breath periods (and therefore breathing frequency and instantaneous minute ventilation), and imprecise location might result in the inspired gas fractions being sampled at an inappropriate time (eg too soon after the genuine end of expiration).

Using the damped mouth pressure signal (as described in the last chapter §2.2.2) has two flaws. Firstly, the detection of the end of a breath is delayed by both the damping and the use of a threshold pressure level. Secondly, any drift in the manometer's zero level during the experiment will skew the ratio of inspiration to expiration (Fig 3.1). A more reliable method for detecting ends of breaths uses inspiratory and expiratory flow, in a method similar to that developed by Lamarra (Patil et al, 1984).

In ideal circumstances either the inspiratory or the expiratory pneumotachograph, but not both, will be registering a flow at any one time. The end of a half breath would thus be simply the time when one flow ceases and the other begins. In practice it is not so simple. Brief fluctuations in the flows, too rapid to be genuine breaths, can be registered as breaths by the program; and noise on the signals and pauses between breaths mean that the pneumotachographs can sometimes both register flows simultaneously. The following method was therefore adopted.

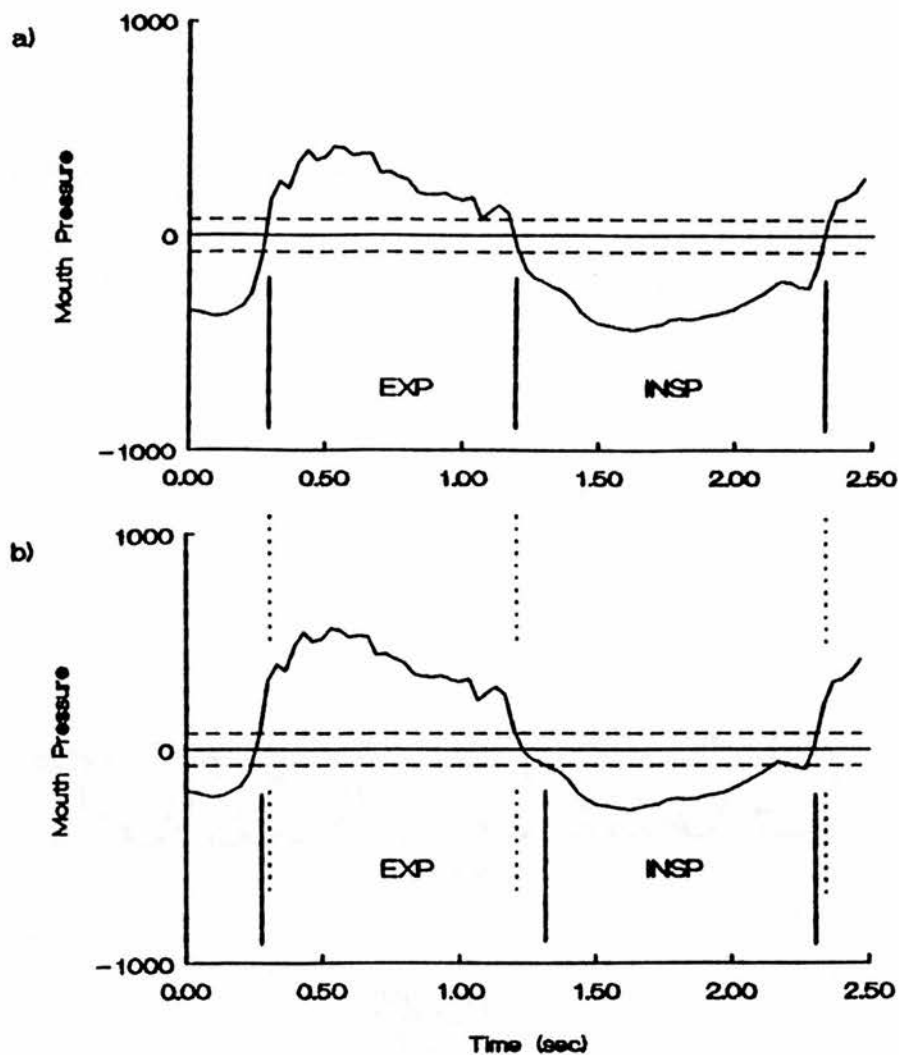


Fig 3.1 The effect of a drift in the manometer zero level on end of breath detection. a) and b) show the same mouth pressure trace but with different zero level. In b) expiration appears to be increased. The broken lines show the threshold levels (somewhat exaggerated).

Each time their output signals were sampled (ie 60 times per second, initially), the inspiratory and expiratory pneumotachographs were defined as being "on" or "off". A pneumotachograph was "on" if the current reading was above a threshold value, the last reading was above $2/3$ of the threshold and the last but one was above $1/3$ of the threshold. Similarly, the pneumotachograph was "off" when three consecutive readings lay below $1/3$, $2/3$ and $1/3$ times the threshold. Inspiration was taken as the time from when the inspiratory pneumotachograph became "on", the expiratory pneumotachograph having reached "off", to when the inspiratory pneumotachograph was "off" AND the expiratory was "on" (Fig 3.2). This logic resulted in pauses (where both flows were "off") being taken as part of the preceding breath phase (eg a pause was end-expiratory rather than early inspiratory). Precedence was given to a pneumotachograph which was "on" already over one which had just become "on" in cases where both flows were "on" together.

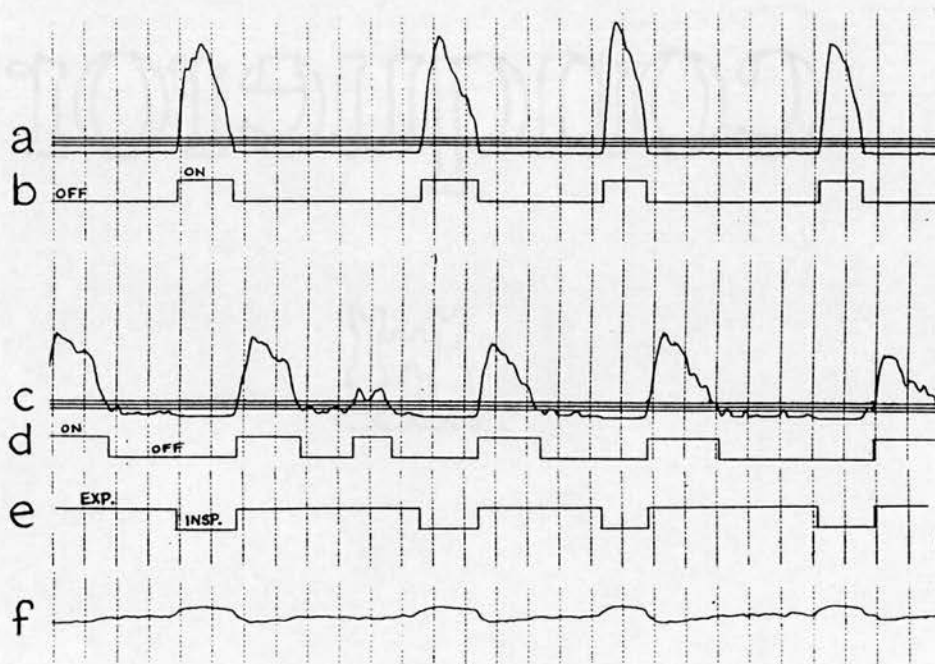


Fig 3.2 End of breath detection. *a and c show inspiration and expiration flow traces respectively, each with the thresholds marked. b and d show the pneumotachograph on/off signal and e the resultant breath signal. f is mouth pressure. The vertical lines are at 1 second intervals.*

There were several advantages to this method. Fluctuations lasting for less than three samples did not affect the decisions; nor did noise lower than the threshold level. A genuine onset would be detected in three readings after passing the $1/3$ threshold level, sooner than if the decision were made only after three consecutive readings above the full threshold. All the signals sampled by the computer, apart from those from the mass spectrometer, were delayed within the program (see Chapter 2, §2.3.3). By reducing the delay for the flow values by the time of three samples, the breath decision was brought into line with the time the flow began to increase towards the $1/3$ threshold level. The exact start of a breath phase was thus pinpointed with increased precision.

The threshold was set by trial and error to about 1 l/min, the minimum consistently effective value. This was about $1/20$ of the maximum flow when breathing at rest, and was sufficiently high to block out noise when using filtered flow signals (see following section). This section of program was checked by outputting a digital signal to a chart recorder with value 1 during inspiration and 0 during expiration, and comparing this with the flow traces (Fig 3.2).

3.2 THE DIGITAL FILTER

On exercise there was a considerable amount of noise on the flow signals, some of it with a regular period (this can be seen in Fig 3.2). The cause of this was uncertain. It may have been due to oscillations of the supporting arm of the valve or to vibrations from the treadmill, but attempts to duplicate the noise using these in isolation failed. Most likely the noise was due to standing waves, set up by the movement of walking, in the column of air from trachea to valve. This hypothesis was supported by the frequency of oscillation varying slightly between individuals, presumably due to variations in length of trachea.

This noise can affect the detection of ends of breath phases. Early attempts using the method just outlined detected "breaths" of extremely short duration due entirely to such fluctuations. It was necessary to reduce such noise on-line if breath detection was to be accurate, and this was most conveniently done in software with a digital filter.

3.2.1 Choice of filter:

A low-pass filter was chosen to have the following characteristics (Fig 3.3):

- Sampling frequency = 30 Hz
- Attenuation < 1 dB for frequencies below 3.6 Hz
- Attenuation > 30 dB for frequencies above 6.4 Hz

The noise to be removed had a frequency of more than 6 Hz and frequencies below this should be attenuated as little as possible. Ideally the difference in frequency between the 1 dB and 30 dB limits should be as small as possible, so that frequencies below the frequency to be filtered out should be attenuated as little as possible. Unfortunately, the narrower this range, the more terms are required in the algebraic expression, and so the resulting calculations require more computing time. A short program was written on a microcomputer to find the number of terms (or "order") required by different specifications, and the specification chosen was the one that had the sharpest possible cut-off for a sixth order filter.

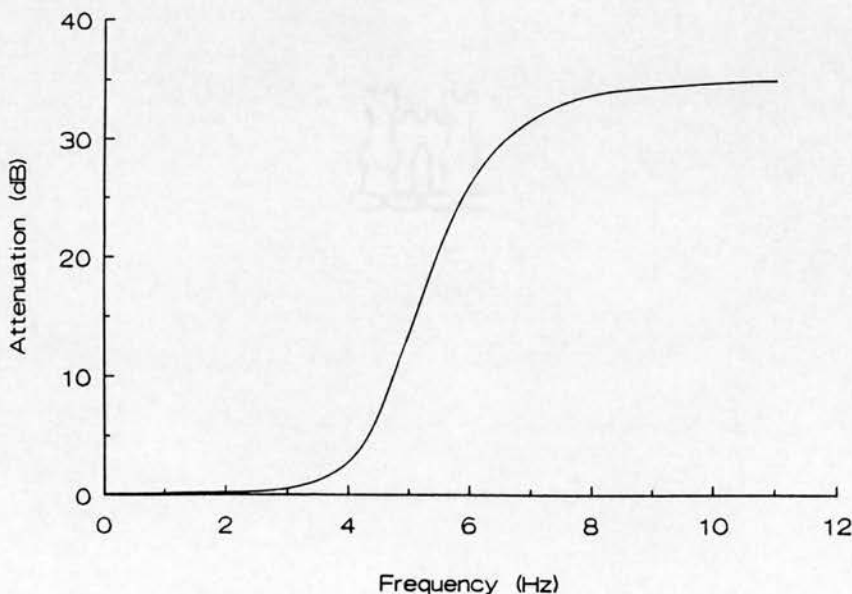


Fig 3.3 The filter function of the chosen Butterworth filter. The plot shows the attenuation in decibels for each frequency.

Initially a more ambitious filter had been planned with a slightly sharper cut-off (3Hz to 6Hz) to work with 60 Hz sampling. This was found to need an eighth order filter, and when the routine was implemented it was found to take 6.5 msec for the two flows during each sample interval, which was slightly more than the time available between samples given the other calculations required. Both the sampling rate and order of the filter were therefore reduced.

3.2.2 Implementation:

Using standard formulae for Butterworth filters (Antoniou¹⁹⁷⁹), the normalised Butterworth filter

$$A = 10\log(1+(\omega^{2n}))$$

(where A is the attenuation in decibels, ω is the frequency in rad/s and n is the order of the filter) was de-normalised (ie stretched) to give the desired frequency response, converted from its analog form to digital using the z-transformation, and manipulated to get the "transfer", or "filter", function in the form

$$H_D(z) = \frac{\sum_{i=0}^N a_i z^{-i}}{1 + \sum_{i=1}^N b_i z^{-i}}$$

Once the values for a_i and b_i had been found, this form of the transfer function could be translated directly into a filtering routine, as z^{-1} was equivalent to T, the sampling interval. The routine therefore multiplied the latest sampled value and those values sampled times T, 2T,... nT ago by the appropriate factors, did likewise with the last output value and those from times 2T, 3T ... nT ago, and summed these values to arrive at the new output. The parameter values had to be specified to double precision or else the filter gradually introduced a bias into the output.

3.2.3 Checking the implementation and finding the delay:

The filtering routine was initially implemented on a microcomputer and was tested thoroughly by simulating various inputs. First, a constant value

of 2 was used as input to check that DC levels were not being altered (which would be a sign of incorrectly entered parameters). The output eventually settled to give a constant reading of 2.0000007, which was of sufficient accuracy. A program was written to input sine waves of different frequencies and unit amplitude and measure the amplitude of the resulting output sine wave once the output had settled (the initial fluctuations in amplitude were due to the sudden change from constant zero to an oscillating input, which excites certain frequencies and results in a process known as "ringing" until those frequencies die away). Filters of this nature introduce a delay between input and output which varies with frequency, and so the program also printed out a value for this delay alongside the attenuation for each input frequency. The plot of attenuation against frequency was exactly as intended, and a delay of $4T$ (ie four sampling intervals) was found to be the average for the pass band.

The routine was copied into the main data acquisition program, and the sampling frequency was reduced to 30Hz to allow time for the filter routine to be called between each sample. The reduction in sampling does not make a significant difference to any of the variables being stored: generally values were averaged over a breath, and little useful information was gained by using 60Hz. An extra time lag of 4 sample intervals was added to the time all values other than flow were stored in buffers (see §2.3.3 for an explanation of why buffering was required) in order to bring these values in line with the filtered flow signals.

3.3 DETECTION OF SWALLOWING

Some subjects, especially those unused to using the mouthpiece, swallow frequently. This interrupts the breathing pattern, usually affecting the tidal volume and duration of the breaths adjacent to the swallow, and this will affect assessment of the respiratory drive if such incidents are not dealt with appropriately. In order to remove these breaths from the data file, the mouth pressure, end-tidal gas and flow signals (recorded on a time-based chart recorder) were visually examined for signs of swallowing. During a swallow, both flow and mouth pressure were zero for an uncharacteristic length of time (ie longer than the pauses between breaths), and frequently end-expiratory CO_2 and tidal volume showed sharp changes. Locating these breaths was a time-consuming process, and the resulting decisions

were inclined to be subjective. It was felt that the process could be automated.

A program was developed with the following aims:

- To provide a method to pinpoint the occurrence of swallowing which works effectively at any rate of breathing or level of exercise. Breathing patterns and the effect of swallowing can vary considerably with the level of exercise.
- To keep the method as rapid and simple as possible, so that it can be carried out on-line.
- To minimise the number of missed swallows, and also the number of mistaken positive identifications.
- To achieve a success rate at least as good as the existing method of detection.

3.3.1 Identifying swallowing:

To determine the criteria for identifying swallowing, nine healthy subjects were asked to breathe on the system for ten minutes, initially at rest, and then walking at a speed comparable with those used in the main studies. They indicated whenever they swallowed and the time was recorded. Inspiratory and expiratory flows, and expired volume (expiratory flow integrated from the start of expiration), were sampled at 60 Hz throughout and stored directly onto a disk, to be plotted after the study. Examples from the traces are shown in Fig 3.4.

Due to the action of the epiglottis, one cannot breathe and swallow at the same time, so during swallowing both flows must be zero (allowing for noise on the signal). In the experiments swallowing was associated with an abnormally long end-expiratory or end-inspiratory pause, or a mid-expiratory or mid-inspiratory pause. Mid-expiratory or mid-inspiratory pauses happened infrequently in normal breathing, and were almost always associated with a swallow. On the other hand, end-inspiratory and especially end-expiratory pauses occurred regularly in the breathing patterns of some of the subjects at rest, and even on exercise in two subjects, making swallows harder to distinguish. Swallowing was

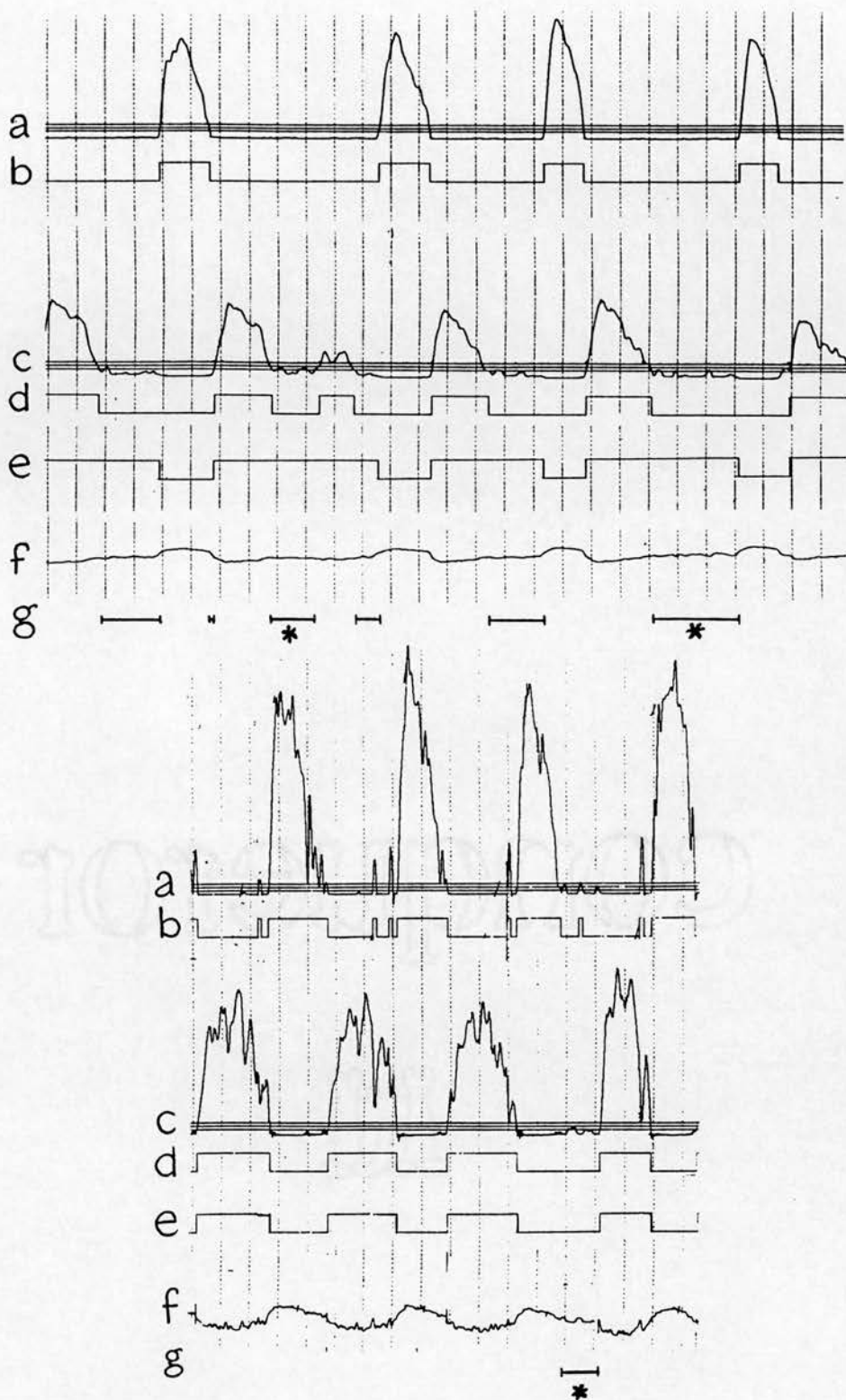


Fig 3.4 Examples of swallows. The upper set of traces is during rest, the lower on exercise. In each case, a and c show inspiratory and expiratory flow, b and d the pneumotachograph on/off signals, e the calculated breath signal and g shows pauses. The asterisks mark a mid-expiratory and two end-expiratory swallows, all identified successfully by the algorithm.

frequently (though not always) followed by a breath of unusually high tidal volume, but similarly large tidal volumes, quite unrelated to swallowing, were also recorded occasionally.

The observations were used to construct algorithms for detecting swallows. These were incorporated into a computer program which read the sampled pneumotachograph data collected on disks exactly as though reading directly from the pneumotachographs. The version with the highest success rate in identifying swallowing in the collected data was copied into the main data acquisition program. This set a marker in the output file with every breath associated with a swallow, and these breaths could later be removed off-line.

3.3.2 The most successful algorithm:

The mean and standard deviation of the duration of the end-expiratory pauses of the latest 8 breaths were calculated, and if the current end-expiratory pause was greater than the mean plus 4 standard deviations the pause was deemed a swallow (and was not used in calculating the mean length for subsequent breaths). A further condition was that the pauses had to be a minimum of 0.5 seconds long: none of the subjects swallowed in less than 0.6 seconds, and this condition prevented the occasional short pause during exercise (when the mean length of pause was frequently zero) to be misidentified as a swallow. A similar approach was applied to end-inspiratory pauses. Mid-expiratory or mid-inspiratory pauses were detected more simply. If a pause of more than 0.5 seconds followed at least 0.5 seconds of one of the flows being "on", and was itself followed by the same flow being "on" again for 0.5 seconds, then a swallow had occurred. The time limits were to prevent transient fluctuations causing false identification of swallows.

These criteria were the most successful of a long trial and error search, varying the number of standard deviations used, the number of breaths used in the averaging, the minimum times for accepting pauses and the use of tidal volume as a criterion. Tidal volume was not found helpful: it was much more variable breath by breath than pause length, and resulted in more false negatives (ie missed swallows).

3.3.3 The success rate

Table 3.1 gives the number of correct, false positive and false negative identifications for each of the nine subjects when the above algorithm was applied to their data. The algorithm was highly successful in 3 subjects (1, 2 and 3) but completely unsuccessful in one (9, who has a very low breathing rate, with long expiratory pauses, even on exercise). Few swallows were missed, except in Subject 6 (whose tendency towards exceptionally long expiratory pauses allowed her to swallow without increasing the pause length).

False positive identifications were the main source of error. These were almost entirely due to the occasional prolonged end-expiratory pause (swallowing in other parts of the cycle being identified with greater accuracy). Such low numbers of false positives do little damage: 3 or 4 normal breaths removed due to mis-identification out of a total of more than 500 will not alter results. However, false negatives can alter results by affecting the breathing pattern.

The data acquisition program incorporating the final algorithm was tested on 3 new subjects at rest and while walking at 2 or 3 different speeds (changing their tidal volume and breathing frequency regularly). The subjects indicated whenever they swallowed and the times were noted. Table 3.2 shows the results for these subjects. Excellent results were achieved for Subject 11, but several swallows were missed (false negatives) in Subjects 10 and 12. In some cases this was because the level of noise occasionally took the flow above the threshold level during pauses, reducing the length of the pause as measured by the computer, and thereby disrupting the algorithm. This was corrected by increasing the threshold.

Two experienced workers (A and B) were asked to locate the swallows in the data of Subject 10 in their usual manner, and their results were compared with those of the computer and the times of actual swallows (Table 3.3). In Table 3.4, the 44 breaths marked by the program or by one of the experimenters have been split up to show the number that each marked which were genuine swallows; and the experimenters and program were then compared in pairs, showing the number of breaths

that each pair both marked or did not mark, and the number that only one of them marked. This shows up the disparity between the two experimenters, who agreed on only 14 of the swallows, whilst between them they marked 22 (ie 19+3) breaths to be removed on which they did not agree. Experimenter A was in closer agreement with the computer program, but marked very many more breaths. With 2 fewer correct swallow identifications but 6 fewer misidentifications, the program scores higher than Experimenter A.

Human operators have very different views of what should be rejected. In terms of detecting genuine swallows, the computer had the highest success rate, but the algorithm did not detect breaths that were suspect in some other way. Defining what "suspect" means is difficult: breaths with rapid noisy fluctuations at low flow tended to be rejected by the two experimenters, as were ones with patterns that were markedly different from the rest. These characteristics can be picked out quickly by a practised operator, but are hard to define precisely enough for a computer program (or even for teaching consistently to another person). The algorithm developed for the program fulfilled all the criteria set out beforehand, but it demonstrated that humans are frequently performing more complex tasks than they had supposed.

Table 3.1 - The success of the algorithm in identifying swallowing in the data on which the algorithm was based.

Subject	Correct	False positive ie not real	False negative ie missed
1.	11	0	0
2.	27	1	0
3.	8	1	0
4.	9	2	0
5.	11	3	0
6.	10	0	3
7.	7	4	0
8.	7	4	1
9.	0	11	2
TOTAL	90	26	6

Table 3.2 - A test of the algorithm on new data (the values in brackets correspond with marked disruption in the breathing pattern).

Subject	Treadmill speed m/s	No. of Breaths	True Positive	True Negative	False Positive	False Negative
10.	Rest	50	3	41	4	2
	1.0	73	8	62	(2)	1
	1.35	102	5	93	0	4
	1.45	58	3	52	0	3
						cont/



11.	Rest	108	8	98	(2)	0
	1.0	46	3	43	0	0
	1.35	71	2	68	0	1
	1.45	74	3	71	0	0
12.	Rest	45	5	39	0	1
	1.0	71	4	64	2	1
	1.45	73	2	68	0	3

Table 3.3 - Comparison between two operators and the computer algorithm of their success in detecting swallowing in the data of Subject 10, showing the algorithm's higher success rate.

Experi- menter	No. of Breaths	True Positive	True Negative	False Positive	False Negative
A.	283	23	243	10	7
B.	283	10	246	7	20
Computer	283	21	249	4	9

(Actual no. of swallows: 30)

Table 3.4 - Tables to show the levels of agreement between the experimenters and algorithm over the 44 breaths which were marked as swallows by one or more (of which only 30 were actual swallows).

"+" = breaths marked as swallows "-" = breaths not marked

		B		Computer	
		+	-	+	-
A.	+	14	19	17	16
	-	3	8	8	3
B.	+			5	12
	-			20	7

Of the 44 breaths:

Experimenter	A.	B.	computer
Total marked	33	17	25
Total unmarked	11	27	19
Total correct	27	17	31

CHAPTER 4

THE EXPERIMENTS

4.1 INTRODUCTION

This chapter describes the experimental methods and results of the study that provided the data on which the later chapters are based.

The aim of the study was to compare the values of hypoxic drive obtained from measurements using three types of stimulus: transient hypoxia, step changes in hypoxia lasting approximately three minutes, and progressive hypoxia. The study has been described in detail in the PhD thesis of Dr. SC De Cort (1989) and so will only be presented briefly here, describing only the results relevant to the remainder of this thesis. In particular, this thesis is concerned only with the responses to step and transient hypoxia, so progressive hypoxia will be given only cursory attention. Before presenting the experimental method, the different types of stimulus are described in more detail, giving their relative merits from the point of view of an engineer testing the response of a system and of a physiologist trying to overcome important practical problems.

4.1.1 Testing a system's response from an engineer's point of view:

A system may be thought of as a "black box": we want to know what is inside the box, but the box cannot be opened, so its contents must be deduced from examining the outputs resulting from different inputs. Much can be deduced simply by feeding various levels of input into the box and measuring the resulting output level for each one: this is the "steady state" method.

Steady state experiments cannot fully describe a system's response as they give no indication of how the system would react to a changing stimulus. This "dynamic" response is often described by breaking it down into the system's responses to all possible frequencies of input sine waves (see §1.4). If the system is linear, knowing the responses to all frequencies is sufficient to calculate the response for ANY pattern of input. Such

knowledge is obtained typically by using standard stimuli, the most frequently used being sine waves, impulses, steps and ramps.

A **sinusoidal input** will generally produce a sinusoidal output, though often with a changed amplitude and phase. By finding the amplitude and phase changes for different frequencies of sine wave, the overall response can be mapped out. This is easily done for, say, an electrical system, but it can be a cumbersome and time consuming process when applied to systems where very low frequencies are important (such as the respiratory system) due to the long collection times required and the difficulties of controlling accurately the shape of the input. Such a method has, however, occasionally been used in ventilatory control experiments (Stoll, 1969; Daubenspeck, 1973; Swanson & Bellville, 1974; Robbins, 1984).

An **impulse** (or "delta function") is, strictly, a sudden sharp stimulus of infinite magnitude and infinitesimal duration. This is equivalent to the sum of sine waves of every frequency all with the same phase. The response to an impulse is therefore a summary of the responses to all frequencies, and is in fact the same as the "convolution function". In practice, an impulse need only contain frequencies up to the highest frequency of interest to be of use, and it therefore has a finite height and width. The greater the height and the narrower the width of an impulse, the higher the maximum frequency it contains.

The integral of the impulse is the perfect **step**, an infinitely rapid change from 0 to 1 at time 0. Again, all frequencies are included in such a stimulus, so the response is a complete summary of the system's response. Frequently a step is easier to simulate in an experiment than an impulse as a single sudden change to the stimulus is sufficient, whereas the impulse requires the stimulus to be switched on and then immediately removed. Steps, like impulses, need only contain frequencies up to the highest of interest. As the level of high frequencies is reduced, the step becomes less sharp: the edges curve and the slope becomes less steep.

Ramps are less useful in the analysis of dynamic responses. The stimulus here increases with constant gradient through the course of the test. High frequencies are not well represented in such an input, so the output from a single ramp does not summarise the system's response. Ramps of varying gradients can be used for examining rate dependent effects, and a ramp

will highlight non-linear effects such as a sudden change in response at a threshold value of stimulus. Generally in respiratory studies, however, the ramp stimulus has been used to obtain approximations to the steady-state response for different levels of stimulus rather than to examine dynamic response. It can take more than 10 minutes to reach steady state when an hypoxic stimulus is administered from a normoxic baseline, so a study requiring several steady state values can be time-consuming and uncomfortable for the subject. After a small increase in stimulus starting from steady state, a close approximation to the new steady state value of response is quickly reached. A series of such small increases in stimulus allows the system to remain in a condition where the response is always near that of the steady state, so a series of values of steady state response may be obtained in a relatively short period. Clearly, the aim here is to eliminate the dynamic aspects of the response, and such studies will not play a major part in this thesis. The concept of more rapidly changing ramp stimuli, however, is used in the series of studies examining the possible effect of rate of change of stimulus on the response (Chapter 6).

In practice it is impossible to obtain an infinitely rapid change in stimulus in any circumstances, and this is especially so in the case of studies of hypoxic ventilatory response. The main sensor of hypoxia, the carotid body, detects P_{O_2} in the blood. For exercising human subjects, the only current ethically acceptable method of manipulating this is by controlling the inspired gas mixture. It takes time for a sudden change in inspired gas mixture to be reflected in the alveolar gas mixture as each inspired tidal volume is mixed with a greater volume of gas already in the lung. Further blurring of the time course of the stimulus occurs as the alveolar gas equilibrates with the blood, so that the carotid body experiences a very much less sharp change in oxygen tension than that applied to the inspired gas.

The rate of change of stimulus at the carotid body depends on the size of the change in inspired oxygen fraction: a change from 21% to 0% O_2 , as in the nitrogen transients, will produce a steeper change at the carotid body than a change from 21% to 12% O_2 , as in the step changes. Additionally, the longer duration of the steps will reveal more clearly any effects in the response which take time to reach their maximum levels, while the shorter duration transients will highlight faster-acting effects.

4.1.2 Physiological and practical considerations:

P_{a,CO_2} : The response of the carotid body to hypoxia is strongly affected by the level of arterial P_{CO_2} . If the hypoxic response is to be measured accurately, P_{a,CO_2} must be kept constant, preferably at the normal value for that subject. An increase in ventilation due to hypoxia will decrease alveolar P_{a,CO_2} as a greater volume of CO_2 will be expired with each breath. P_{a,CO_2} will in turn be reduced, and this will in turn inhibit the response of the carotid body. This leads to the experimentally awkward situation where the response to a prolonged hypoxic stimulus causes attenuation of that response, making it difficult to obtain an accurate assessment of the purely hypoxic component of the response.

Transient stimuli: The problem of reduced P_{a,CO_2} is avoided to a considerable extent in using transient stimuli. It takes of the order of 5 seconds, the exact time depending on the subject's cardiac output, for the blood with reduced P_{a,O_2} to reach the carotid body after the onset of the stimulus (Jain et al, 1972). During this time the ventilation is at its normal level as the stimulus has not reached the detector. Typically on the 3rd breath following the switch to nitrogen the ventilation increases, and by the end of expiration P_{a,CO_2} will be somewhat reduced: again it takes several seconds for this hypocapnic blood to reach the carotid body so for perhaps two breaths the blood at the carotid body is hypoxic without being hypocapnic. The blood becomes progressively more hypocapnic over the next few breaths, and this will reduce the ventilatory response. If the stimulus is two breaths of nitrogen, the hypoxia will be reduced anyway at a similar time to the start of hypocapnia, so the "on-stroke" of the response, the increase in ventilation, should not be affected by the hypocapnia, although the "off-stroke" will be more rapid. If the stimulus is 3 breaths of nitrogen, the carotid body will experience a period where the hypocapnia is beginning to increase while the hypoxia is still increasing, so the on-stroke may be affected. Maximum hypocapnia, however, does not reach the carotid body until after the hypoxic blood has passed. The exact timing and number of breaths depends on several factors, including the cardiac output and breath periods.

A drawback of using transient stimuli is that the response is measured over a very few breaths. There is a degree of randomness in the value of

\dot{V}_E for a given breath and this can be a sizeable proportion of the increase in \dot{V}_E caused by the hypoxia. It is necessary to average the responses to several transients to reduce the effect the randomness has on the assessment of hypoxic drive. Implicit in this is the assumption that each transient involves the same sequence and timing of events although, as has been stated above, these are affected by the durations of breaths and thus indirectly by \dot{V}_E . On the positive side, transients are easy to administer (although it is hard to control the size of stimulus), and cause little discomfort to the subject.

Step stimuli: The longer duration of step stimuli means that CO_2 has to be added to the inspired gas mixture once \dot{V}_E has increased in order to maintain P_{a,CO_2} at its normal level. The standard method for achieving this is to add sufficient CO_2 to the inspired gas to keep end tidal P_{a,CO_2} at its pre-hypoxia value (see §4.2.5 below).

Randomness in \dot{V}_E has a much lower effect on assessments of hypoxic drive made with steps than with those from transients due to the greater number of breaths involved. Steps can however prove uncomfortable for some subjects.

Prolonged hypoxia acting centrally tends to reduce the ventilatory response (see §1.2.5). This becomes significant typically 5-10 minutes after the start of hypoxia (Holton & Wood, 1965; Weiskopf & Gabel, 1975). Steps lasting less than 3 minutes should therefore avoid problems associated with this effect.

Progressive reduction of P_{O_2} : During progressive hypoxia studies the level of hypoxia is gradually increased over 10 minutes. As with steps, CO_2 must be added to the inspired gas to keep P_{a,CO_2} constant. The long duration involved means that central effects are likely to affect the level of response, and it is for this reason that progressive hypoxia does not play a role in the remaining chapters of this thesis, in which it is assumed that the only factor affecting \dot{V}_E is the increase in hypoxia.

4.2 METHODS

4.2.1 General:

The studies were all approved by the Lothian Health Board Ethics of Medical Research Sub-committee for Medicine and Clinical Oncology.

Subjects fasted for at least 4 hours prior to a study, and were allowed to drink only water, as caffeine and the digestion of food affect the respiratory drive (Lakshminarayan et al, 1978; Zwillich et al, 1977).

The subjects walked at a steady pace on a level treadmill, the speed being adjusted so that the subject's oxygen uptake (\dot{V}_{O_2}) was approximately 1.0 l/min. An initial period of at least 10 minutes was allowed for gas exchange to reach steady state (§2.2.9). All control of the experiment was carried out behind the subject so that the inspired gas concentration could be changed without the subject's knowledge. The gas mixtures were switched during expiration so that the change could not be felt, and the subject wore a personal stereo playing music to block out any aural clues.

4.2.2 Protocol:

Ten healthy subjects, four female and six male, all with normal lung function, each underwent a practice session and two study sessions. The practice session was used to determine the number of breaths of nitrogen needed to take the subject's S_{a,O_2} to about 80% during a transient and the speed of the treadmill at which the subject's \dot{V}_{O_2} was approximately 1.0 l/min, as well as to allow the subject to acclimatise to the equipment. One of the study sessions involved transient and progressive hypoxia; the other, the one of interest for this thesis, was composed of:

- 6 transients followed by 10 minutes rest and then 4 steps (15% O_2 , 12% O_2 , 15% O_2 , 12% O_2) separated by 5 minutes breathing room air.

4.2.3 Transient Hypoxia:

Once steady-state conditions had been attained, the inspired gas mixture was switched from room air to 100% nitrogen. The subject took 1, 2 or 3 breaths of nitrogen to give a fall in S_{a,O_2} to approximately 80%, and the gas

mixture was then changed back to room air. Sixty breaths were allowed for recovery before the process was repeated.

The fall in S_{a,O_2} for a given number of breaths of nitrogen depends on the duration and the ratio of tidal volume to residual lung capacity for each of the breaths. These factors vary from breath to breath, and so the fall in S_{a,O_2} is likely to change from one transient to another. Studies therefore generally consisted of several transients (frequently 6) to obtain sufficient number with the desired fall in S_{a,O_2} .

4.2.4 Step Hypoxia:

The inspired gas was switched from room air to the gas mixture from the rotameters set to either 12% or 15% oxygen, the balance being nitrogen. As ventilation increased, CO_2 was added to the inspired mixture to keep P_{ET,CO_2} constant (see below). After 3 minutes the valve was switched back to room air and a recovery period of at least 5 minutes was allowed before the next hypoxic step. If the subject's S_{a,O_2} fell below 80% during the step, or the subject showed signs of discomfort (for instance due to a high \dot{V}_E), the inspired gas was immediately switched to room air.

4.2.5 Control of P_{a,CO_2} :

A widely used method for controlling P_{a,CO_2} during prolonged hypoxia is to add CO_2 to the inspired gas mixture in sufficient quantities to keep end-tidal P_{CO_2} (P_{ET,CO_2}) constant. As this process is referred to in Chapter 11 and Appendix A, it will be described in a little more detail.

The concentration of CO_2 in the expired gas was displayed on a chart recorder throughout the study. During hypoxia the experimenter watched this trace and attempted to keep the breath-by-breath peak value of expired CO_2 concentration (which is directly related to P_{ET,CO_2}) as near to the average peak value prior to the step as possible by adjusting the flow of CO_2 . The flow required depended upon the tidal volume of each individual breath, and frequent small adjustments were made as the peak varied. The circuit used did not allow P_{ET,CO_2} to be kept constant at the end of the step: the subject was simply returned to breathing room air, even though \dot{V}_E might still be high for several breaths. There was therefore a period of hypocapnia following each step (Fig 4.1).

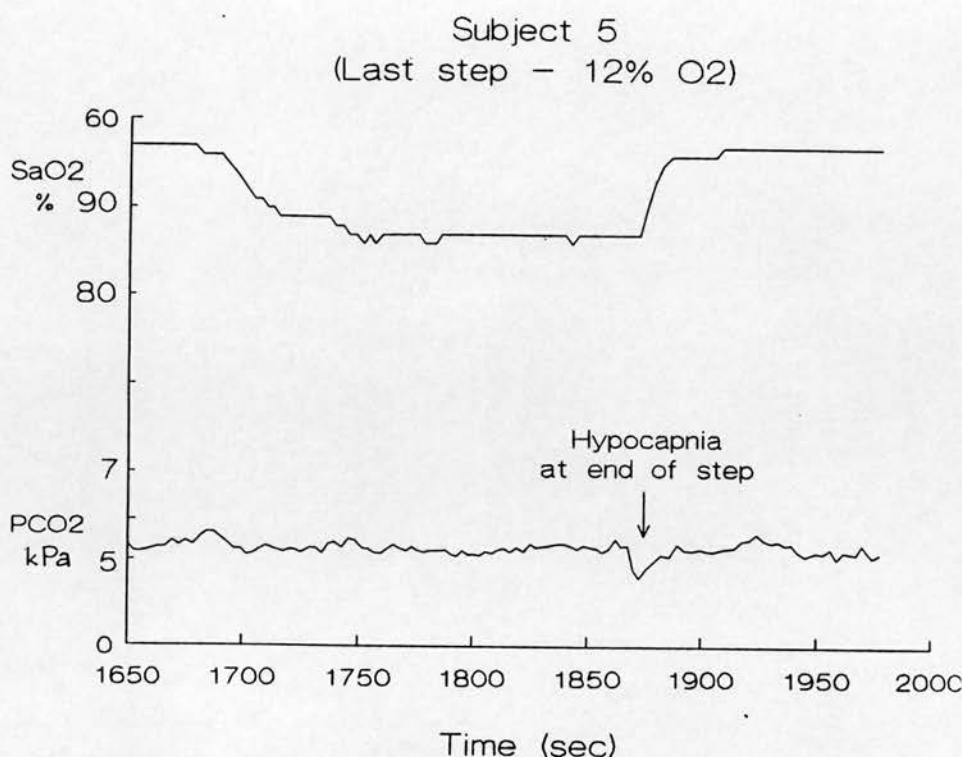


Fig 4.1 *The control of P_{a,CO_2}* The lower trace shows how P_{a,CO_2} was kept at its pre-stimulus level during hypoxia until the end of the step when it was allowed to fall.

4.3 RESULTS

\dot{V}_E and the gas exchange variables (\dot{V}_{O_2} and \dot{V}_{CO_2}) during steady-state exercise prior to the transients were not significantly different (Friedman's analysis of variance) from those measured before the steps (Table 4.1), indicating that the levels of exercise and metabolism were the approximately the same. The mean baseline P_{ET,CO_2} also did not vary significantly between these two periods, and the mean and standard deviation of P_{ET,CO_2} during the steps (when P_{ET,CO_2} was being controlled by the addition of inspired CO₂) did not differ significantly from those of the baseline (Table 4.2). P_{ET,CO_2} did, however, fall significantly below the baseline mean level during the transients.

The subjects responses to the different stimuli are given in Fig 4.2(a-j). In eight of the subjects the overall response to step hypoxia was higher than to transient hypoxia for the same fall in S_{a,O_2} , whilst in the other two

(Subjects 7 and 9) this relationship was reversed. This is confirmed by the slopes of \dot{V}_E vs S_{a,O_2} plotted for the pooled transients and the pooled steps (Table 4.3). The values for the transients were obtained by correcting each for the delay in S_{a,O_2} caused by the oximeter, separating out the 3 points preceding the stimulus and all the points until S_{a,O_2} reached its minimum value, pooling this set of points for all 6 transients and finding the linear regression. The slopes calculated for the steps were from the linear regression of the pooled data from all 4 steps, using the 10 points preceding the stimulus and all points up to the end of hypoxia (ie again not using the period where S_{a,O_2} was returning to baseline).

The difference in response to steps and transients was not small: in 6 of the subjects (2, 3, 4, 5, 8 and 10) the \dot{V}_E - S_{a,O_2} slope for steps was more than double that for transients. In the most extreme case (Subject 3, Fig 4.2c), the difference was more than six-fold. The response of this subject to transients was typically an increase in \dot{V}_E of 10 l/min for a fall in S_{a,O_2} of 22%, whilst during one step the increase in \dot{V}_E reached 40 l/min for a fall in S_{a,O_2} of 10%. Subjects 7 and 9 (Figs 4.2g & i) showed a lower response to steps than to transients. For example, Subject 7's increase in \dot{V}_E for a transient fall in S_{a,O_2} of 18% was 30 l/min, whilst the increase for a step fall of 16% was only about 12 l/min.

Fig 4.2 (Next few pages) Plots of \dot{V}_E , S_{a,O_2} and PET_{CO_2} for all ten subjects. The vertical scales are the same for all subjects.

Fig 4.2a Subject 1

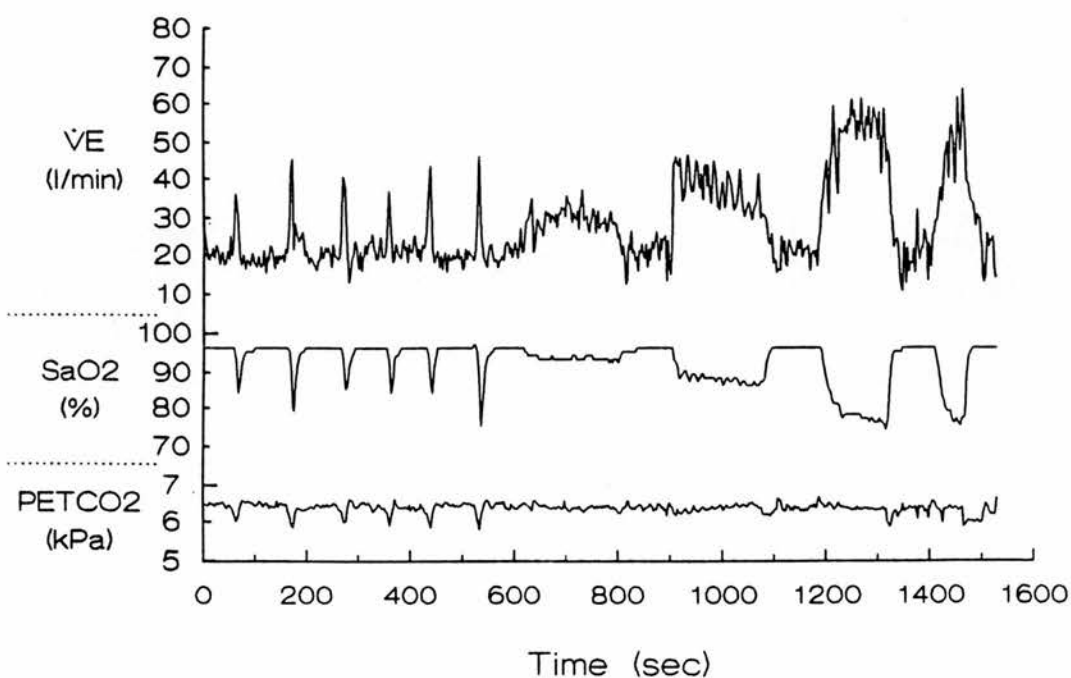


Fig 4.2b Subject 2

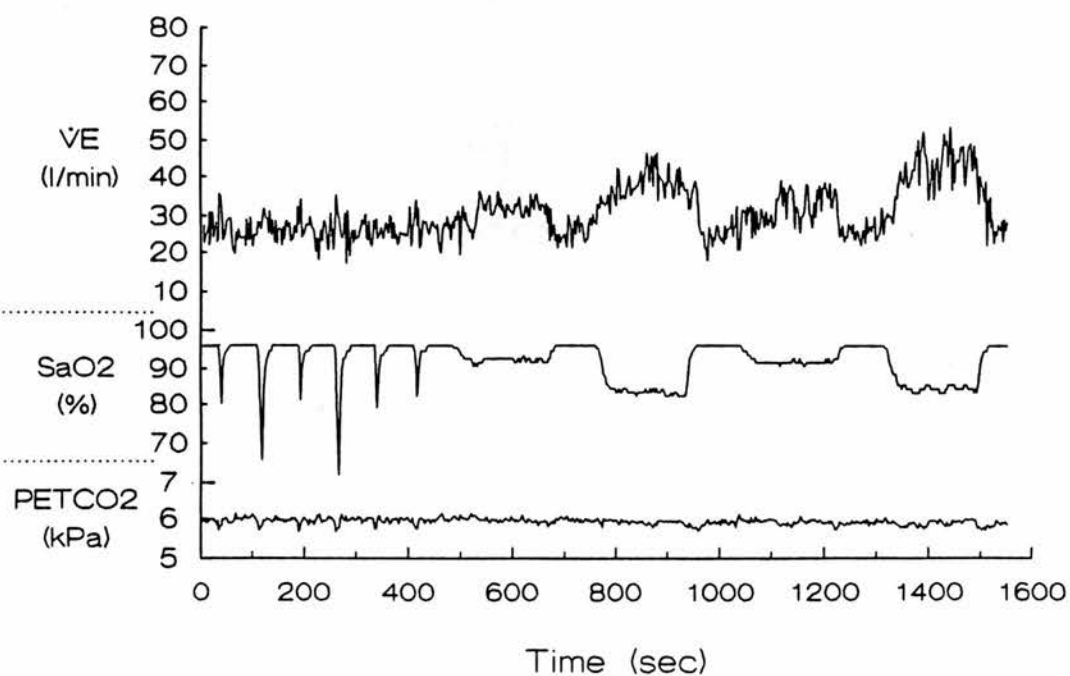


Fig 4.2c Subject 3

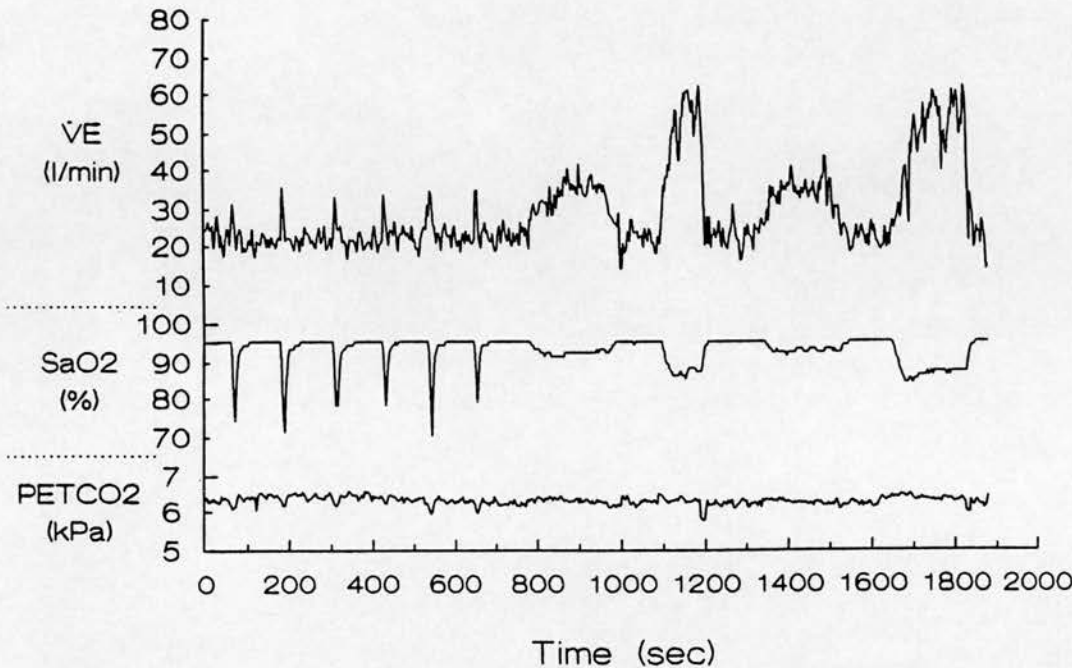


Fig 4.2d Subject 4

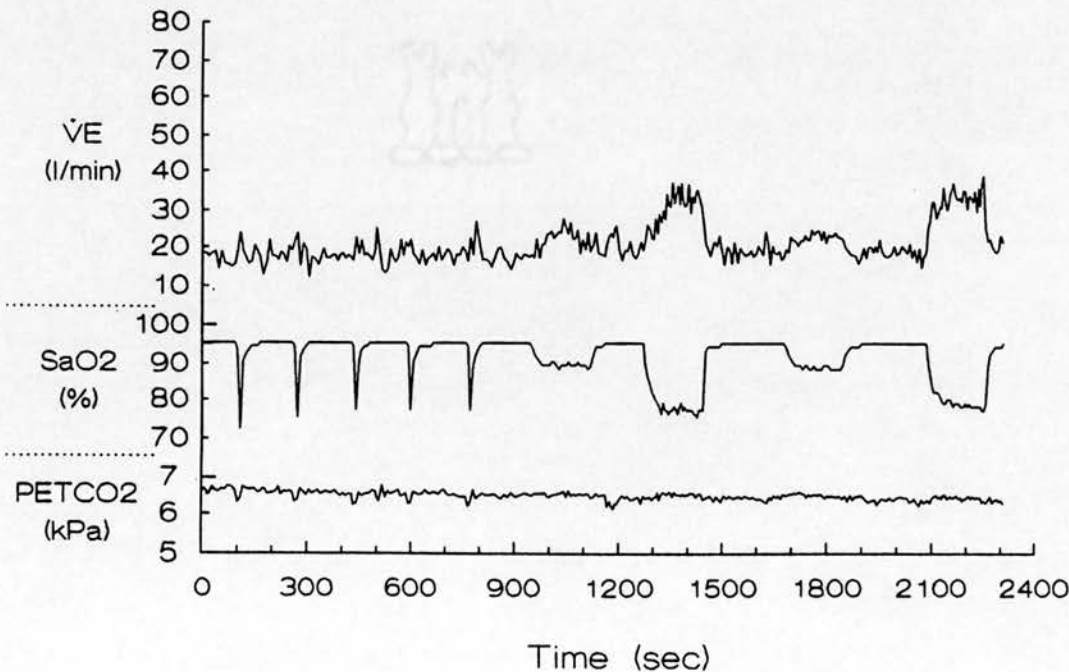


Fig 4.2e Subject 5

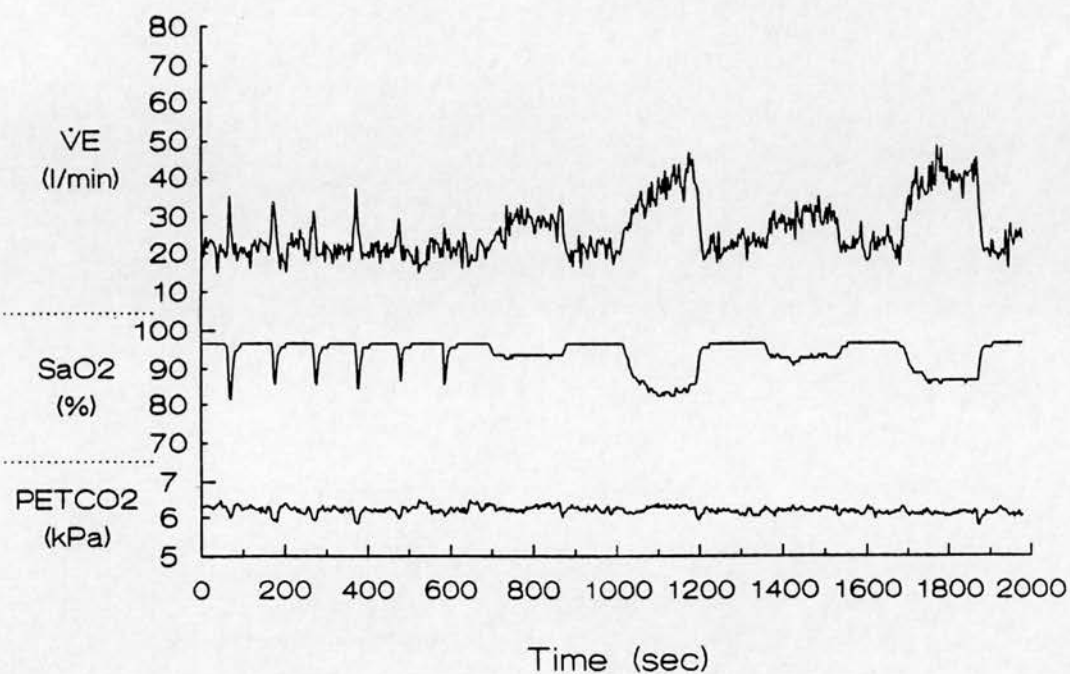


Fig 4.2f Subject 6

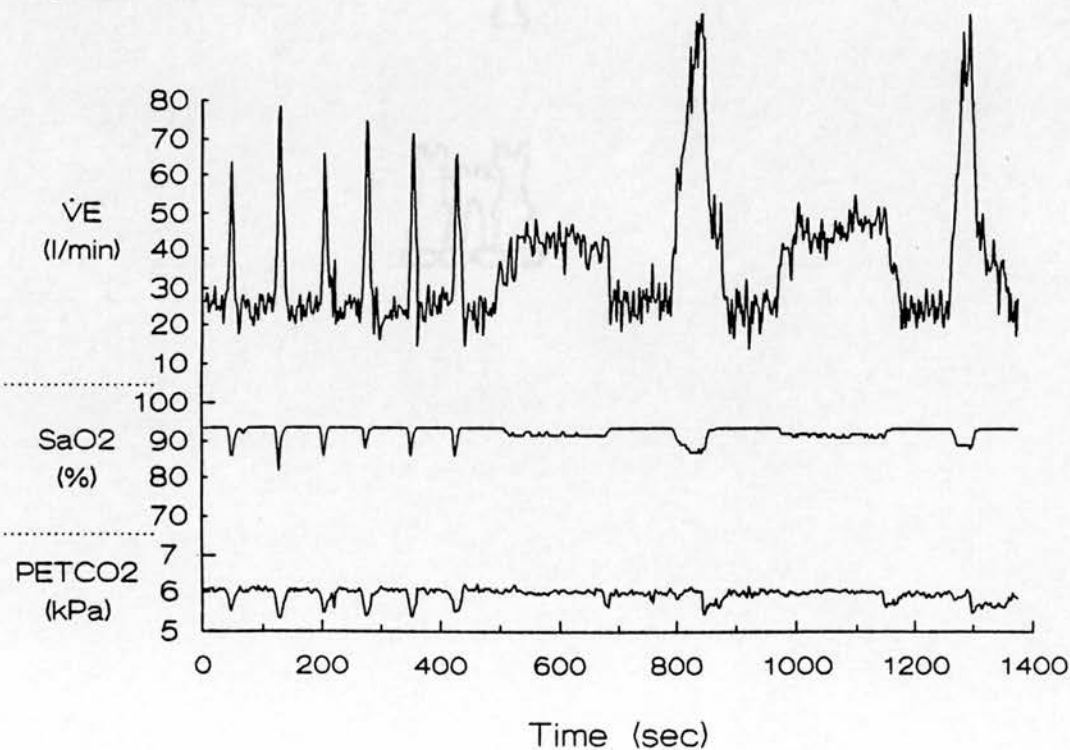


Fig 4.2g Subject 7

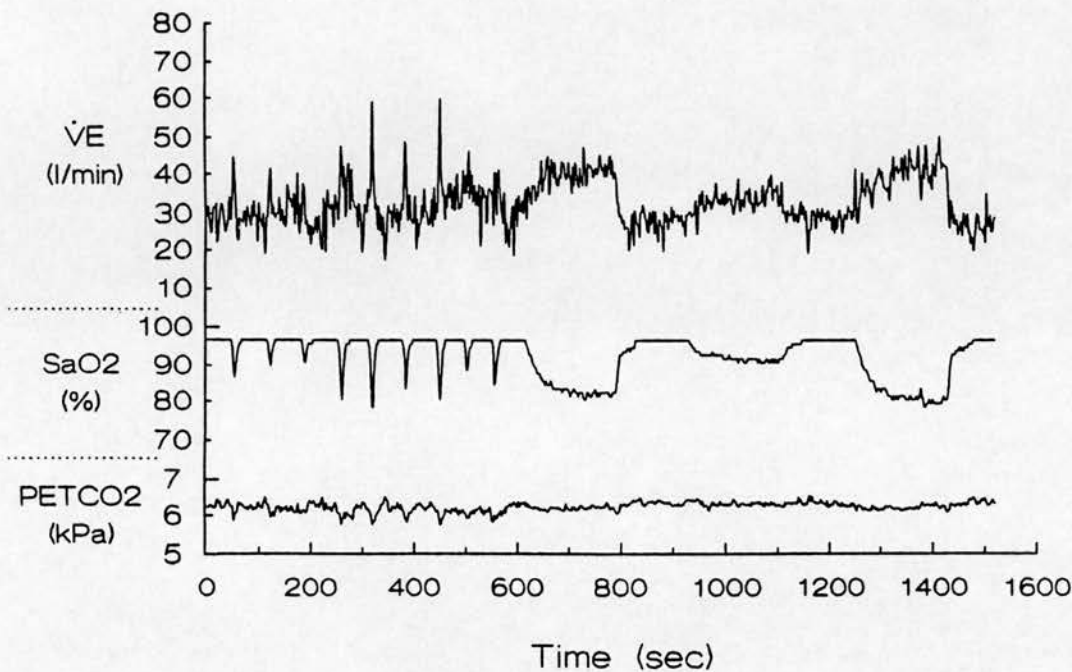


Fig 4.2h Subject 8

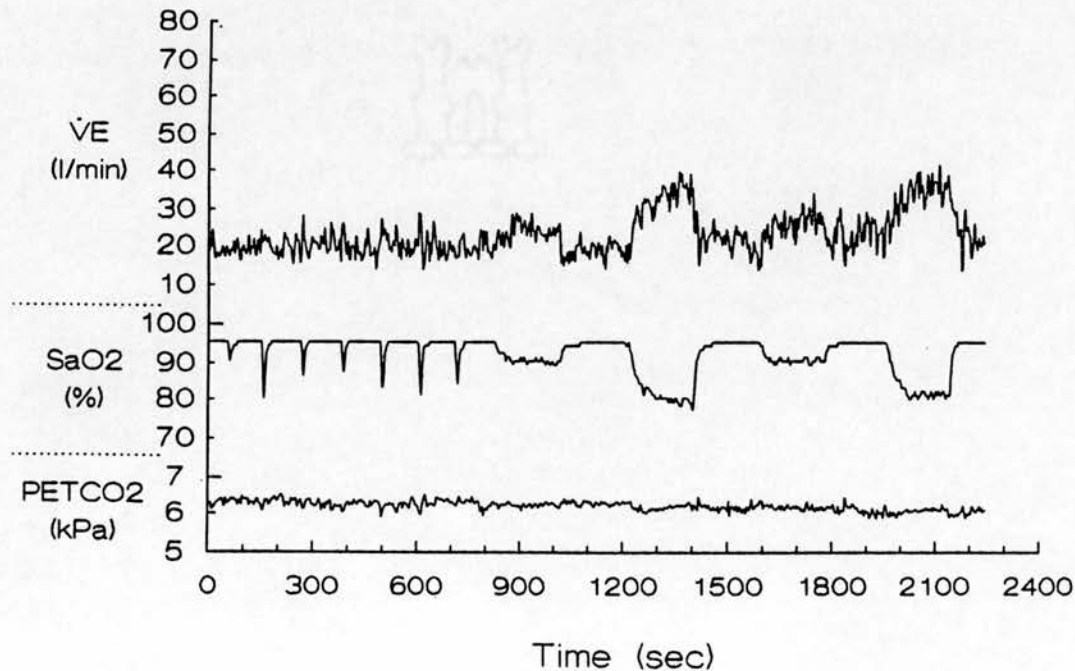


Fig 4.2i Subject 9

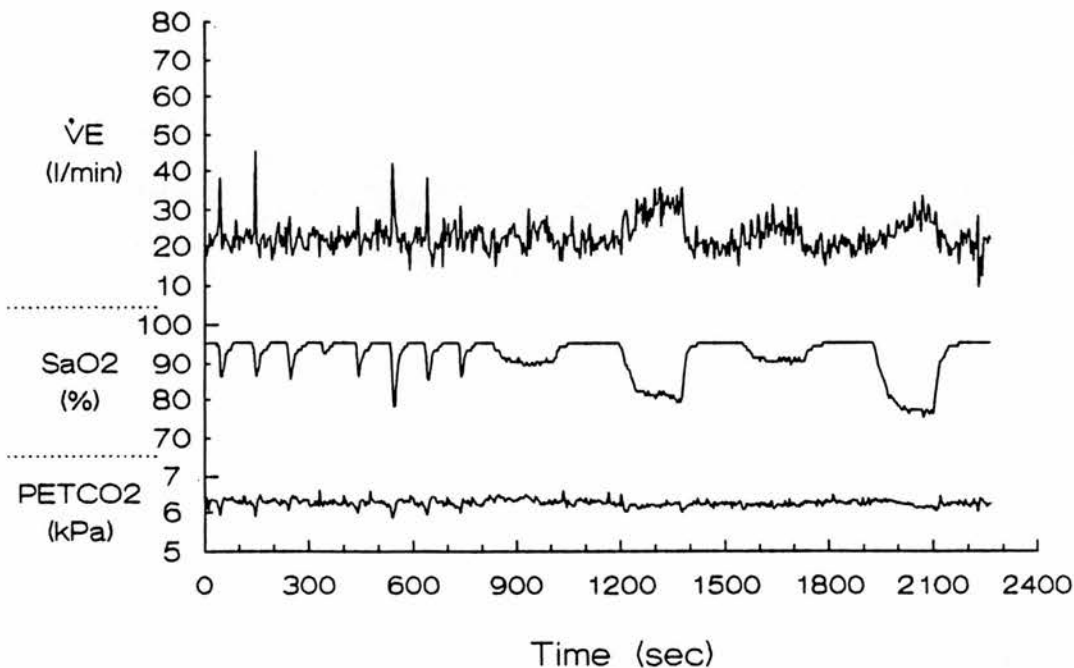


Fig 4.2j Subject 10

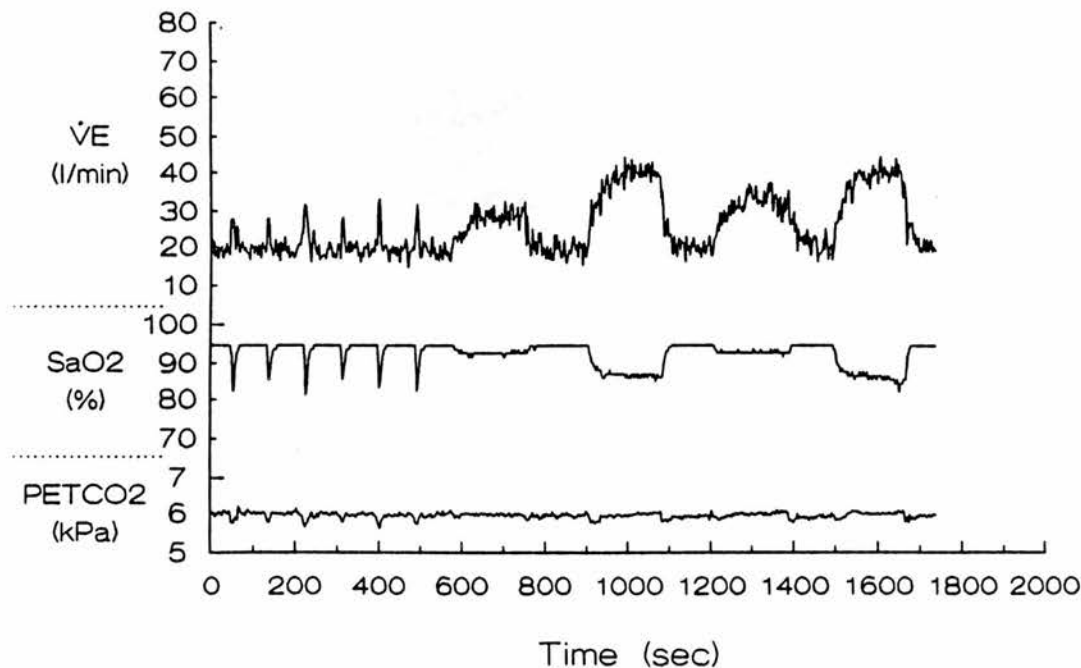


Table 4.1 - Gas exchange variables during steady state exercise prior to the transient and the step hypoxia.

	Transient			Step		
Subject	\dot{V}_{O_2}	\dot{V}_{CO_2}	\dot{V}_E	\dot{V}_{O_2}	\dot{V}_{CO_2}	\dot{V}_E
1	0.98	0.84	22.92	0.93	0.80	23.12
2	0.90	0.84	26.80	0.93	0.88	27.09
3	1.04	0.84	22.51	1.14	0.97	25.19
4	1.01	0.84	19.43	0.99	0.83	19.24
5	0.95	0.80	22.24	0.95	0.78	22.56
6	1.04	0.90	27.86	1.07	0.92	28.07
7	0.96	0.94	29.53	0.99	1.01	28.42
8	0.89	0.78	20.56	0.96	0.83	22.00
9	0.93	0.84	24.02	0.95	0.81	21.56
10	0.79	0.69	20.61	0.78	0.66	21.06

\dot{V}_{O_2} and \dot{V}_{CO_2} are expressed as l/min STPD; \dot{V}_E is expressed as l/min BTPS.

Table 4.2 - Baseline end tidal PCO₂ during the normoxic periods of the transient and step sections of the study, and during the steps and the transients themselves.

P _{ET,CO₂} (kPa)				
Subject	Baseline		During Stimulus	
	Step	Transient	Step	Transient
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
1	5.88±0.13	5.93±0.13	5.82±0.16	5.61±0.39
2	5.01±0.12	5.03±0.09	4.94±0.14	4.87±0.22
3	5.61±0.15	5.84±0.17	5.68±0.14	5.26±0.28
4	5.95±0.19	6.19±0.16	5.46±0.19	6.06±0.27
5	5.44±0.22	5.49±0.16	5.46±0.19	5.51±0.16
6	5.09±0.13	5.16±0.08	5.08±0.12	4.60±0.53
7	5.69±0.17	5.46±0.17	5.54±0.14	5.31±0.34
8	5.35±0.22	5.66±0.18	5.38±0.19	5.51±0.24
9	5.60±0.13	5.62±0.15	5.53±0.18	5.41±0.23
10	4.99±0.08	5.12±0.11	5.00±0.12	4.82±0.24

P_{ET,CO₂} is expressed as mean±SD of the pooled data from:

Column *a.* 20 breaths preceding each step stimulus,
 b. 10 breaths preceding each transient stimulus,
 c. all breaths during steps before return to room air,
 d. all breaths during the on-phase of transients (see
 section 1.3.3).

Table 4.3 - Hypoxic Ventilatory Drive during Transient and Step Hypoxia.

Subject	$\dot{V}_E/S_{a,O_2}$ (l/min/%)	
	Transient	Step
1	-1.40	-1.71
2	-0.40	-1.05
3	-0.47	-3.26
4	-0.22	-0.80
5	-0.72	-1.53
6	-5.27	-9.78
7	-1.03	-0.70
8	-0.43	-0.96
9	-0.91	-0.44
10	-0.78	-1.70

(See text for details of calculations)

4.4 DISCUSSION

The remainder of this thesis is largely an examination of why 8 out of the 10 subjects should show a greater response to step changes in S_{a,O_2} than to transient changes.

Why should there be this difference in response to stimuli that were administered during the same study session at the same levels of baseline \dot{V}_E , \dot{V}_{CO_2} and \dot{V}_{O_2} ? One obvious difference between the administration of transient and step hypoxia was that P_{ET,CO_2} was controlled in the latter but not in the former. This, however, is unlikely to be the full explanation for the difference in response. In §4.2.3 above, it was argued that the hypocapnia resulting from the increase in \dot{V}_E would not affect carotid body response until after a period during which arterial blood at the carotid body was hypoxic but normocapnic. In her PhD thesis, Dr De Cort (1989) examined the effect of adding CO_2 to the inspired gas mixture during transients, and found no statistical difference between the response to these transients and those without CO_2 in the same subjects during the same session.

An alternative explanation of the differences in response to steps and transients, at least for the eight subjects whose response to steps was greater than to transients, was that the effect was simply due to a slow time constant, and that the low response to transients was simply because there was insufficient time for the response to develop to its maximum potential before the end of the stimulus. It is to this hypothesis that the next chapter is devoted.

CHAPTER 5

FOURIER ANALYSIS OF THE DATA

5.1 INTRODUCTION

Eight out of ten subjects in the first study showed what appeared to be a greater response to an hypoxic step than to a transient stimulus. A possible explanation of this is that the subjects' responses to transients and steps were essentially the same, but because the steps were of longer duration the response had more time to develop and so reached higher values of \dot{V}_E . The difference may simply be due to the system having a long **time constant**. This chapter is an examination of this hypothesis.

The main difficulty in comparing steps and transients is that the shapes of the stimuli are quite different, the steps having a lower rate of fall of S_{a,O_2} as well as a longer duration. A widely-used mathematical tool for overcoming such problems and, in effect, standardising the shape of the stimulus, is to find the "**transfer function**" (explained in §1.4) relating output to input for the system. For a "**linear**" system this function will be the same whatever the original shape of stimulus, and it provides a complete description of the relationship between input and output for such a system.

This chapter describes how the ideas of Fourier analysis discussed in Chapter 1 were applied to the system relating \dot{V}_E to S_{a,O_2} .

5.2 APPLICATION OF FOURIER ANALYSIS TO THE \dot{V}_E - S_{a,O_2} SYSTEM

The system being considered will be defined as that which acts upon a given time course of S_{a,O_2} to produce an output which is \dot{V}_E . What happens in between is treated as a "black box", ie the exact mechanisms involved are ignored completely and only the relationship between input and output considered. This relationship, assuming the system is not changing and that all other possibly significant factors are kept constant by

the experimental method, may be represented by a transfer function which should be independent of the time course of S_{a,O_2} , providing that the system is linear.

In such conditions, it would be expected that for a given subject the transfer function derived from step change stimuli would be identical to the transfer function derived from transient stimuli. If this were found to be the case then the difference between responses to steps and transients in that individual would be due to mathematical rather than physiological causes, ie physiologically the system was acting identically on the two types of stimuli, but the outputs appeared different because of the considerable differences in shape of input. If, however, the transfer functions were not found to be identical, then we could conclude that one or more assumptions were faulty, ie that other factors were affecting the response, that the system was acting in a physiologically different way to the two shapes of input, or that the system was not linear.

Finding the transfer function representing the ventilatory response to stimuli is not a new idea, but previously this has usually been for the response to changes in P_{a,CO_2} . The method used was to vary the inspired P_{CO_2} sinusoidally at various frequencies, finding the amplitude and phase shift of the resulting oscillations in \dot{V}_E for each frequency (Stoll, 1969; Daubenspeck, 1973). Swanson and Bellville (1974) refined the method so as to force sinusoidal variations in end-tidal P_{CO_2} (and therefore, presumably, in P_{a,CO_2} , so that the transfer functions calculated were not affected by gas mixing in the lung). Robbins (1984) applied the method to oscillations in P_{ET,O_2} , and found that at higher frequencies the system became less linear.

The method we adopted for finding the transfer function relied more on mathematics than experimental technique, using Fourier transforms to avoid the need to collect more data from the subjects, and to allow the question of the difference between the response to steps and transients to be addressed directly.

5.3 METHOD

The process of finding the transfer functions required several steps:

1. Interpolation of the data to a regular sampling interval.
2. Averaging all the steps and all the transients.
3. Finding the transforms for \dot{V}_E and S_{a,O_2} .
4. Dividing the transform of \dot{V}_E by that of S_{a,O_2} .
5. Deletion of points in the transfer function for frequencies that are not sufficiently well represented to be providing useful information.
6. Display of the transfer function as amplitude & phase.

5.3.1 Interpolation:

The values of S_{a,O_2} and \dot{V}_E from the experiments are each an average over a breath. To allow a discrete Fourier transform the sampling period must be regular, so values for half second intervals were derived from the breath-by-breath data by interpolation using the method of Lagrange (Margenau & Murphy, 1956). This uses the following formula (in its 3rd order form):

$$y = \frac{(x-x_2)(x-x_3)(x-x_4)}{(x_1-x_2)(x_1-x_3)(x_1-x_4)} y_1 + \frac{(x-x_1)(x-x_3)(x-x_4)}{(x_2-x_1)(x_2-x_3)(x_2-x_4)} y_2 \\ + \frac{(x-x_1)(x-x_2)(x-x_4)}{(x_3-x_1)(x_3-x_2)(x_3-x_4)} y_3 + \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_4-x_1)(x_4-x_2)(x_4-x_3)} y_4$$

where y is the interpolated value at x , and $(x_1, y_1), (x_2, y_2)$ etc. are existing points on either side of x .

An example of interpolation using this formula is given in Fig 5.1. The interpolated points follow the original points closely, although sharp rises are preceded by slight dips (as Lagrange designed the formula to give curves with continuous derivatives, ie without sudden changes in curvature). These slight deviations are unimportant in the current context. They give curves similar to those obtained by filtering out very high frequencies, and information about the high frequencies would be lost anyway as part of the discrete Fourier transformation (see §1.4.4).

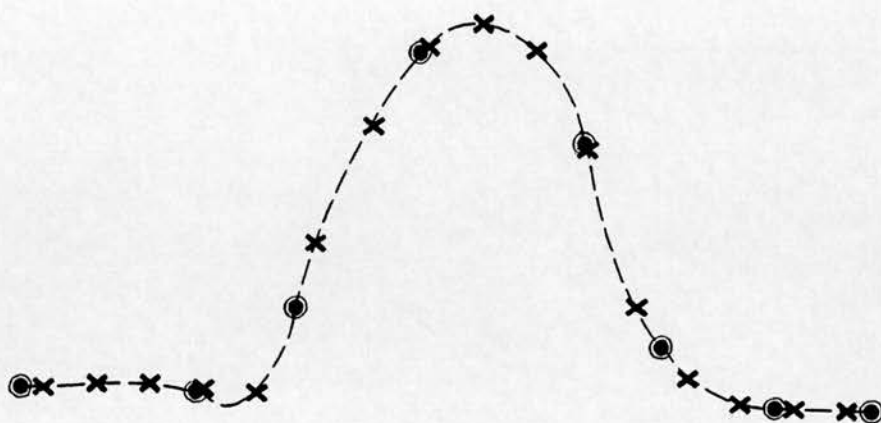


Fig 5.1 Example of interpolated data. Circles show the original, irregularly-spaced data points, crosses show evenly-spaced interpolated points.

5.3.2 Averaging several waveforms:

There is a considerable amount of random noise in the \dot{V}_E waveform, and this will be translated into noise on the transfer function or convolution function. The standard way to diminish the noise on the transfer function would be to average transfer functions derived from several experiments: the random elements then cancel each other out. One property of the Fourier transform is that if F is the transform of f , and G of g , then the transform of $(f+g)$ is $(F+G)$. Thus the mean of two transforms is the same as the transform of the mean of the two functions. This is convenient as it is quicker to find the mean of several waveforms and transform this rather than perform several Fourier transforms and then average.

The mean time course of six transient stimuli (in a single subject) was found by adding together the interpolated traces of 20 breaths before and 40 following the start of hypoxia, aligned according to the time that the switch to the hypoxic gas mixture was made. Each resultant value was divided by the number of points contributing to that value, ie by 6 for most of the points, but by lower numbers near the ends of the traces where

the different lengths of breaths left some traces longer or shorter than others. Similarly 4 step stimuli were averaged for the individual. Again 20 breaths before and 40 after the start of hypoxia were taken.

5.3.3 The transform:

A highly efficient algorithm for performing a discrete Fourier transform using a computer is the **Fast Fourier Transform (FFT)** (Press et al, 1986). By restricting the number of points in the input to a power of two it is possible to cut down considerably the number of individual calculations needed (by a factor of about 30 for 256 points). This method is therefore much faster than the standard method, and was chosen for the program, though this meant that all but 256 points (512 in subjects with low respiratory frequency) were ignored in the analysis. This is equivalent to 128 (or 256) seconds or about 30 breaths.

5.3.4 Finding the transfer function:

For each individual the FFT was performed on the mean interpolated S_{a,O_2} and \dot{V}_E traces for the steps and for the transients. The resulting transforms are a series of complex numbers (ie numbers with a real and an imaginary part). The transfer function was found by dividing each value in the FFT of the \dot{V}_E trace by the corresponding value in the FFT of the S_{a,O_2} using "complex" division. This gave a further series of complex numbers. To find the amplitude and phase changes for each frequency, the following formulae were applied to each complex value z :

$$\text{Amplitude} = \sqrt{\text{Real}(z)^2 + \text{Imag}(z)^2}$$

$$\text{Phase} = \arctan(\text{Imag}(z)/\text{Real}(z))$$

5.3.5 Deletion of points:

Although the change in gas mixture is sudden for both types of stimuli, the change in S_{a,O_2} is much more gradual due to the dilution of the hypoxic gas by residual air in the lungs. The smoother shape of the S_{a,O_2} trace means that higher frequencies in the input are present with only very small amplitudes, and the inherent noise in the signal, along with the imposition by the ear oximeter of digitisation on what should be a

continuous variable, mean that the signal to noise ratio for these high frequencies can be very low. Finding the transfer function involves a division, and this again will decrease the signal to noise ratio so that the transfer function is an unreliable description of the effects on these higher frequencies.

Such misleading points were removed by zeroing all points in the transfer function for frequencies above the value where the FFT of either S_{a,O_2} or \dot{V}_E first fell beneath a cutoff level. This cutoff level was chosen to be where the amplitude fell to 1/1000 of the 0 Hz amplitude, ie where $\text{ampl}(z) < \text{ampl}(z_0)/1000$. At this level the amplitude is close to the resolution of the analogue to digital converter used in sampling the signals being used.

5.3.6 Plotting the results:

The clearest plots were found to be of the transfer function displayed in a manner similar to that used in the design of filters. It is usual to show such curves as attenuation of amplitude measured in decibels (dB) plotted against frequency (see, for example, Fig 3.3). Attenuation is calculated as follows:

$$A = 10 \times \log_{10}(\text{ampl}(z)/\text{ampl}(z_0)) .$$

The reference amplitude (z_0) in electrical filtering is that of the input signal. Using this plot is not entirely appropriate with the \dot{V}_E - S_{a,O_2} system, as, unlike an electrical filter, the input and output are quite different quantities, measured in different units. The 0 Hz "d.c." component is changed by the system from about 97% S_{a,O_2} to about 30 l/min \dot{V}_E . To prevent changes in this relationship between baseline levels from distorting the graphs, the 0 Hz level was used as z_0 and all other values were therefore scaled accordingly. This would result in all other values in the range of interest being negative, so relative amplitude rather than attenuation was plotted (one is just the negative of the other).

With each graph of amplitude there is a corresponding plot of the change of phase (measured in degrees) for each frequency. Because \dot{V}_E rises as S_{a,O_2} falls, the phase changes will basically be centred round 180 degrees: if all phase changes were exactly 180 degrees then \dot{V}_E would mirror S_{a,O_2} exactly. Time delays in the system will show up as phase

changes affecting all frequencies. The phase plots can, however, be difficult to interpret, and care must be taken not to assign too much significance to phase changes at frequencies where the amplitude is low.

5.3.7 Reconvolution:

To show the effects of the transfer functions, the reverse of the deconvolution process described above was carried out. A **standard input** consisting of a square wave fall in S_{a,O_2} of 10% for 10 seconds was convolved with each convolution function (ie the Fourier transform of the standard input was multiplied by each transfer function and the result inverse transformed). This results in a trace of \dot{V}_E . Such traces are directly comparable with each other, unlike the \dot{V}_E traces obtained with different shapes of input.

Differences in transfer function will of necessity show up as differences in the reconvolved trace, and strictly speaking no new information is gained from reconvolving. These traces are, however, more directly comprehensible than transfer functions to the unpractised observer.

5.4 RESULTS

For each of the ten subjects studied, the transfer function was derived separately from the transient and the step change data and these two functions were then compared. Plots of the amplitude and phase change of the transfer functions for each subject are given in Fig 5.2.

It is clear, for some of the individuals at least (1,3,4,6,8 & 10), that the transfer function derived from the step data is not identical to that derived from the transient data. Quantifying the statistical significance of this difference is not possible. There is no reason to suppose that the points should lie on a straight line, nor can it be assumed that points obtained from a series of identical experiments should be normally distributed about a mean point. Strict comparisons of gradients and other parametric tests are therefore ruled out. However, some of the transfer functions give the impression of being straight line relationships within the range plotted, and using this idea makes the **qualitative** differences easier to discuss. Although gradients of the lines will be referred to, it must be remembered that this is an entirely arbitrary method of description.

The most noticeable difference between the two sets of transfer functions is that often the gradient of one is greater than that of the other. This is particularly clear in the amplitude plots for Subjects 4, 6, 8 and especially 9, but close examination shows this to be the case for Subjects 1 and 5 as well. The phase plot for Subject 1 also shows different gradients. The most striking thing about Subject 3's amplitude plot is that the step amplitudes are 4 to 5 times greater than those for the transients. This was the subject with the greatest difference between the response to steps and that to transients. Subjects 2 and 7, on the other hand, showed very close agreement between the step and transient transfer functions.

The corresponding "reconvolved" traces for Subject 4 (Fig 5.3) show that a different magnitude of change in \dot{V}_E results (for a standard input) from the two transfer functions, emphasising their difference.

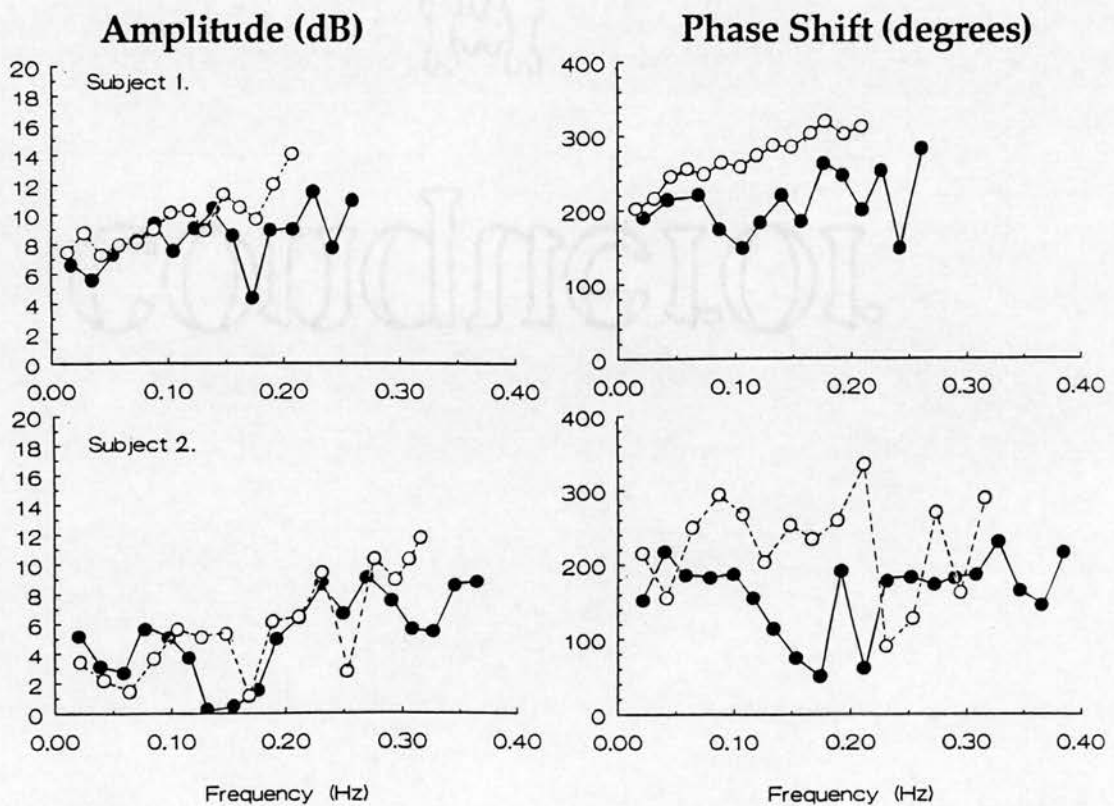


Fig 5.2 (Continued over next few pages) *Amplitude and phase shift plots of transfer functions. Those derived from steps are marked with filled circles, those from transients with open circles.*

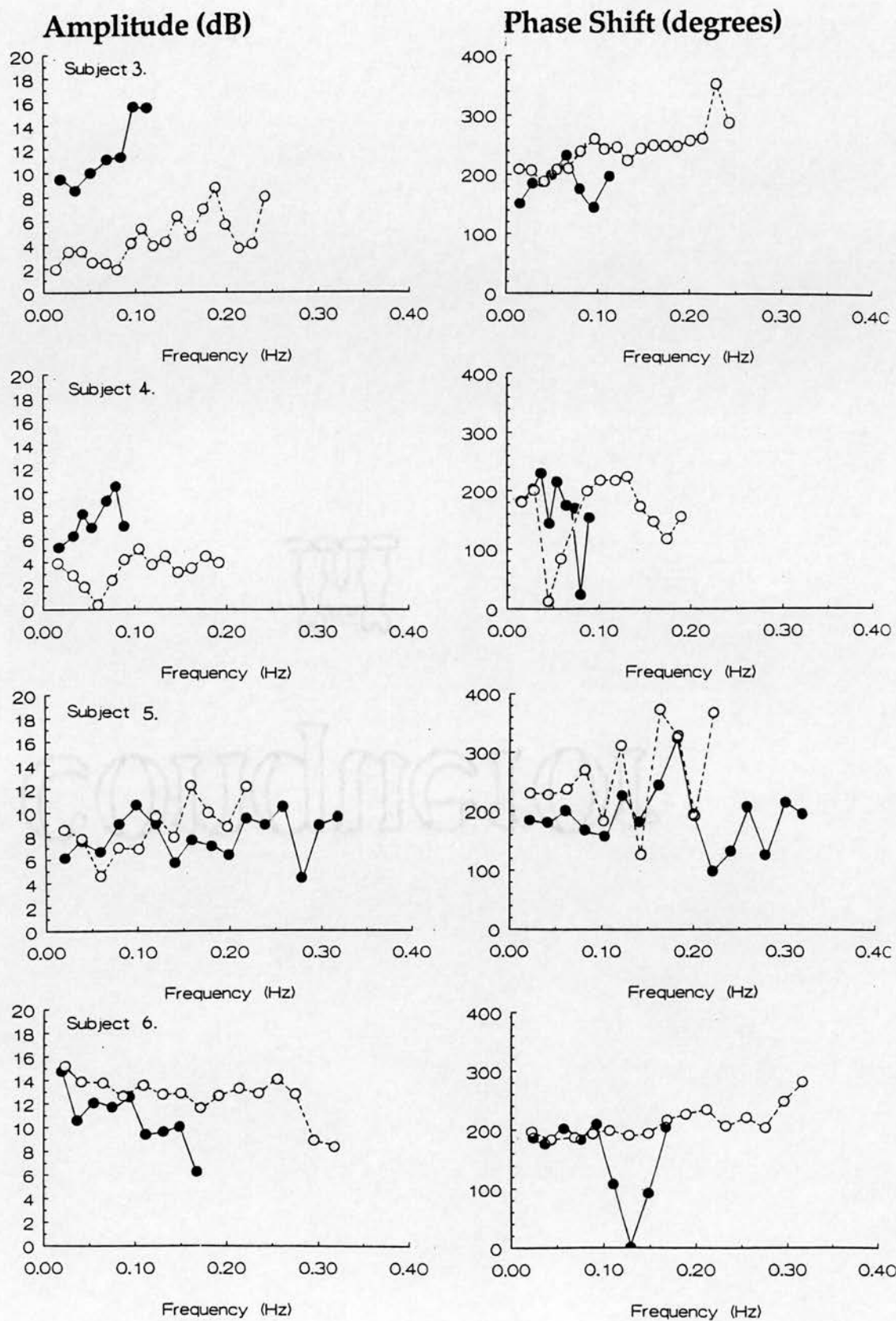


Fig 5.2 continued.

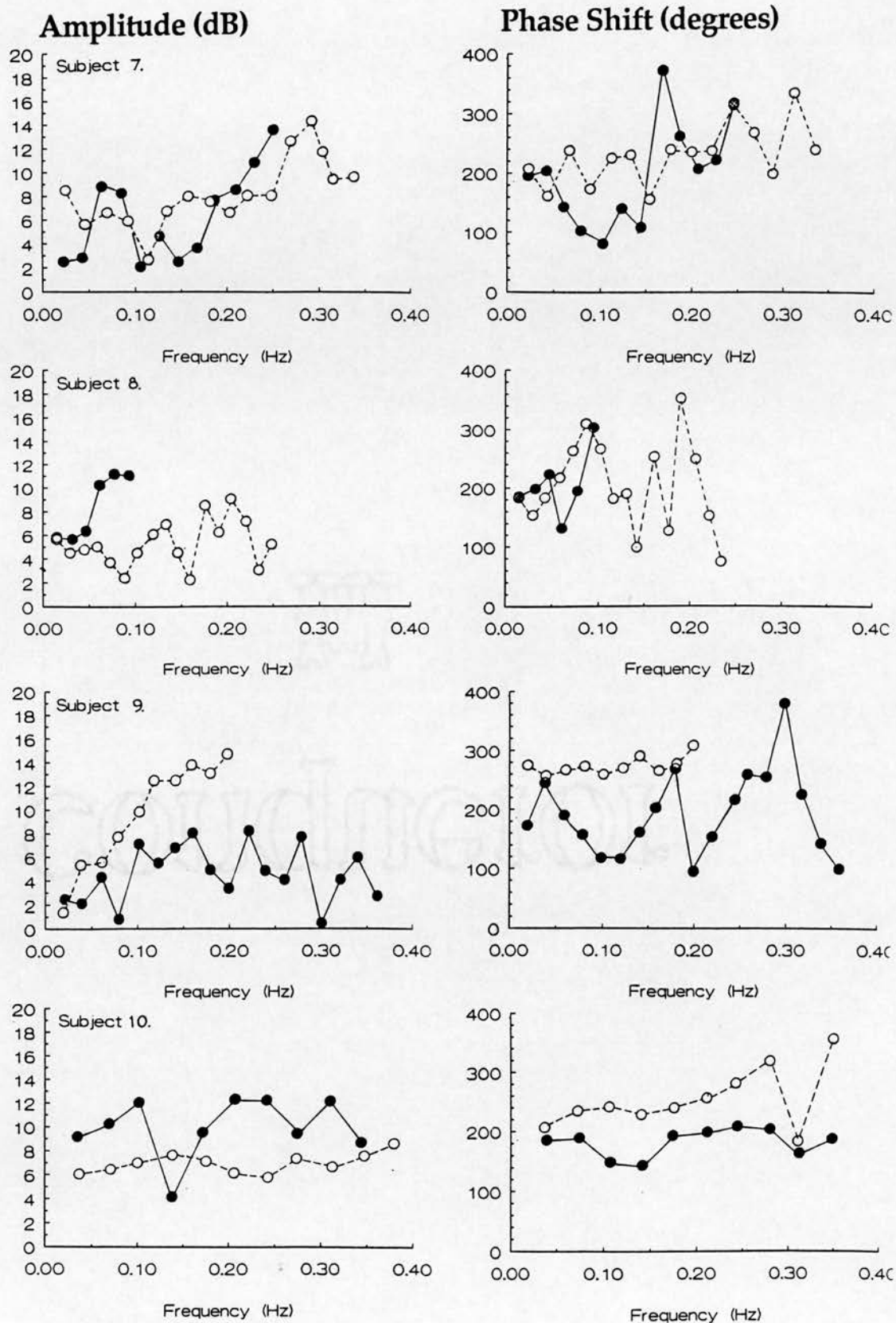


Fig 5.2 continued

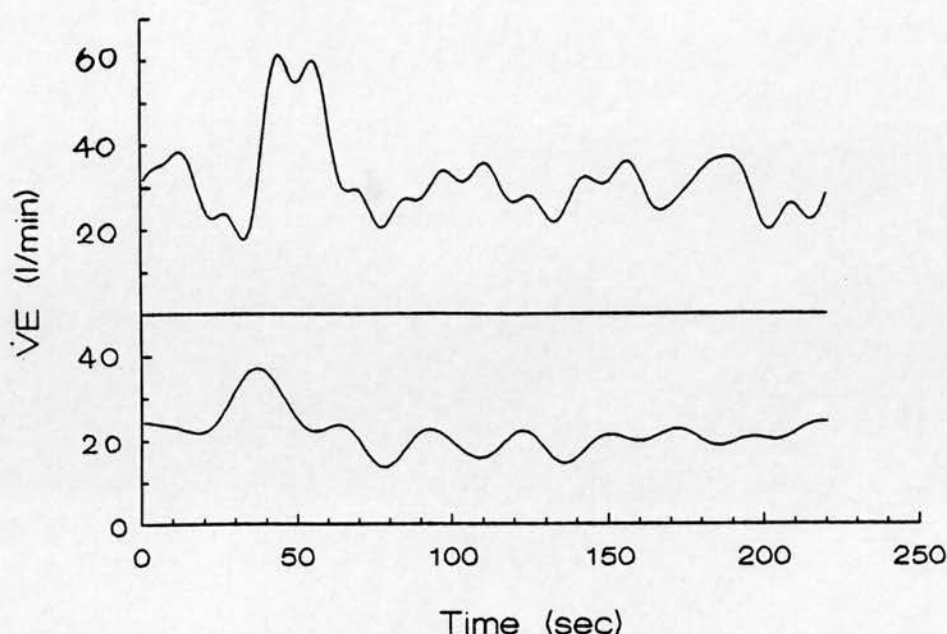


Fig 5.3 Reconvolution. The upper curve was obtained by convolving the step-derived convolution function with a standard stimulus; the lower curve was obtained using the same stimulus and the transient-derived convolution function.

5.5 DISCUSSION

Subject to the following assumptions, the transfer and convolution functions for any given subject should be constant, whatever the shape of stimulus used to derive them. The assumptions are:

1. That no factors other than the fall in oxygen level as measured by S_{a,O_2} were acting as stimuli.
2. That the physiological response of the body was independent of the type of stimulus.
3. That the system is mathematically linear.

These assumptions overlap considerably. For example, if 1. or 2. is false, the system under consideration will necessarily appear non-linear, although the converse is not necessarily the case.

In some subjects at least, the transfer functions, and therefore the convolution functions, were different when derived for a given individual using two types of stimulus. Thus one or more of the assumptions must be wrong. It is only necessary for a few subjects to display this variation in transfer function for such a conclusion to be reached - it is not logically necessary for it to happen in all cases - which is why little attempt has been made to define exactly for which subjects the functions differ.

Assumption 1. - only hypoxia acts as stimulus: Great care was taken in the design and execution of the studies to keep constant all possible stimuli to ventilation other than hypoxia. The chapter on experimental protocols (Chapter 4) should be referred to for details.

Assumption 2. - the physiological response is the same to the two types of stimulus: For this not to be the case, the subject's body would need some way of distinguishing the sort of stimulus it was receiving rather than simply reacting to the fall in oxygen level. It is very unlikely that the differences were due **directly** to the gas mixtures involved (pure nitrogen for the transients and a mixture of nitrogen, oxygen and CO₂ for the steps), but it is quite possible that the **duration** of the stimulus in the case of steps gave time for different mechanisms in the body to affect \dot{V}_E . Another possibility, examined in the next chapter, is that the **rate of onset** of stimulus led to a difference in response.

Assumption 3. - the system is linear: This has been shown to be false. If assumptions 1. or 2. are false then the system we are dealing with is non-linear. If they are both true, then the system must still be non-linear to give the differences in convolution function. **The conclusion must therefore be that the \dot{V}_E - S_{a,O_2} system is NON-LINEAR.**

One consequence of this assumption is that, because the whole process of convolution and deconvolution assumes linearity, the transfer functions obtained are not true transfer functions and are not directly interpretable as such. Little is to be gained from further examination of these functions; instead, experimental and other means must be sought to uncover the nature of the non-linearity in the system. These will be covered in the following chapters.

CHAPTER 6

INVESTIGATING THE EFFECT OF RATE OF CHANGE OF STIMULUS ON RESPONSE

6.1 INTRODUCTION

In the last chapter it was established that the difference seen in responses to transient and step change hypoxia was not simply a manifestation of a slow time constant in a linear system, but was a consequence of some sort of non-linearity in the system. This non-linearity must be dependent on something that is different between the two types of stimulus or it would not have shown up in these experiments. One such difference between the transient stimuli and the step changes is the rate of fall of S_{a,O_2} , which is more rapid in the former.

The hypothesis that the magnitude of response is related to the rate of onset of stimulus was investigated by finding the responses to step changes in hypoxia that began with different rates of fall of S_{a,O_2} . The only difference between these steps was in the first few breaths, so any systematic variation in response would demonstrate that rate of change of stimulus was an important factor. A further interesting line of approach adopted was to continue the analysis of the last chapter using Fourier techniques to see whether the transfer functions changed smoothly between those derived from steps and those derived from transients, or whether the change was sudden at a particular threshold rate of fall of S_{a,O_2} .

6.2 METHODS

The technique described in Chapter 4 for giving 12% O_2 steps in hypoxia during exercise was used with one difference: between switching from room air to the 12% O_2 gas mixture, one or two breaths of either 1% O_2 in nitrogen or 6% O_2 in nitrogen were administered to the subject from a Douglas bag. Following this breath (or breaths) the step proceeded exactly as before with CO_2 being added to keep end tidal P_{CO_2} at its pre-stimulus

level. A practice session prior to the day of the study was used to decide whether one or two breaths from the bag was appropriate for the subject, the aim being to produce an S_{a,O_2} trace that fell rapidly to about the level reached by breathing 12% O_2 without a noticeable overshoot.

A study consisted of 4 types of stimulus:

1. 12% O_2 step,
2. 6% O_2 followed by 12% O_2 ,
3. 1% O_2 followed by 12% O_2 and
4. N_2 transients.

These gave 4 different rates of fall of S_{a,O_2} (in ascending order as listed) and were administered in random order to 5 subjects.

The \dot{V}_E and S_{a,O_2} traces for each pair of steps were interpolated and averaged, and the "transfer functions" were derived as described in the last chapter.

6.3 RESULTS

Full results are detailed in the PhD thesis of Dr JE Hill (1990). She found no significant difference for each subject in baseline \dot{V}_E , \dot{V}_{O_2} , \dot{V}_{CO_2} and P_{ET,CO_2} between the steps. The slopes of the fall in S_{a,O_2} were significantly different, yet the mean \dot{V}_E over the last 30 seconds of each step for an individual did not differ significantly.

Graphs of \dot{V}_E and S_{a,O_2} are plotted in Fig 6.1. No obvious difference in response could be seen between the steps of different gradients, nor was any highlighted by taking the gradients of plots of \dot{V}_E against S_{a,O_2} . Fig 6.2 shows sample transfer functions (for Subject 10) and again no systematic difference could be seen between the steps, though the functions for the transients were different in the same way as found with the earlier series of experiments.

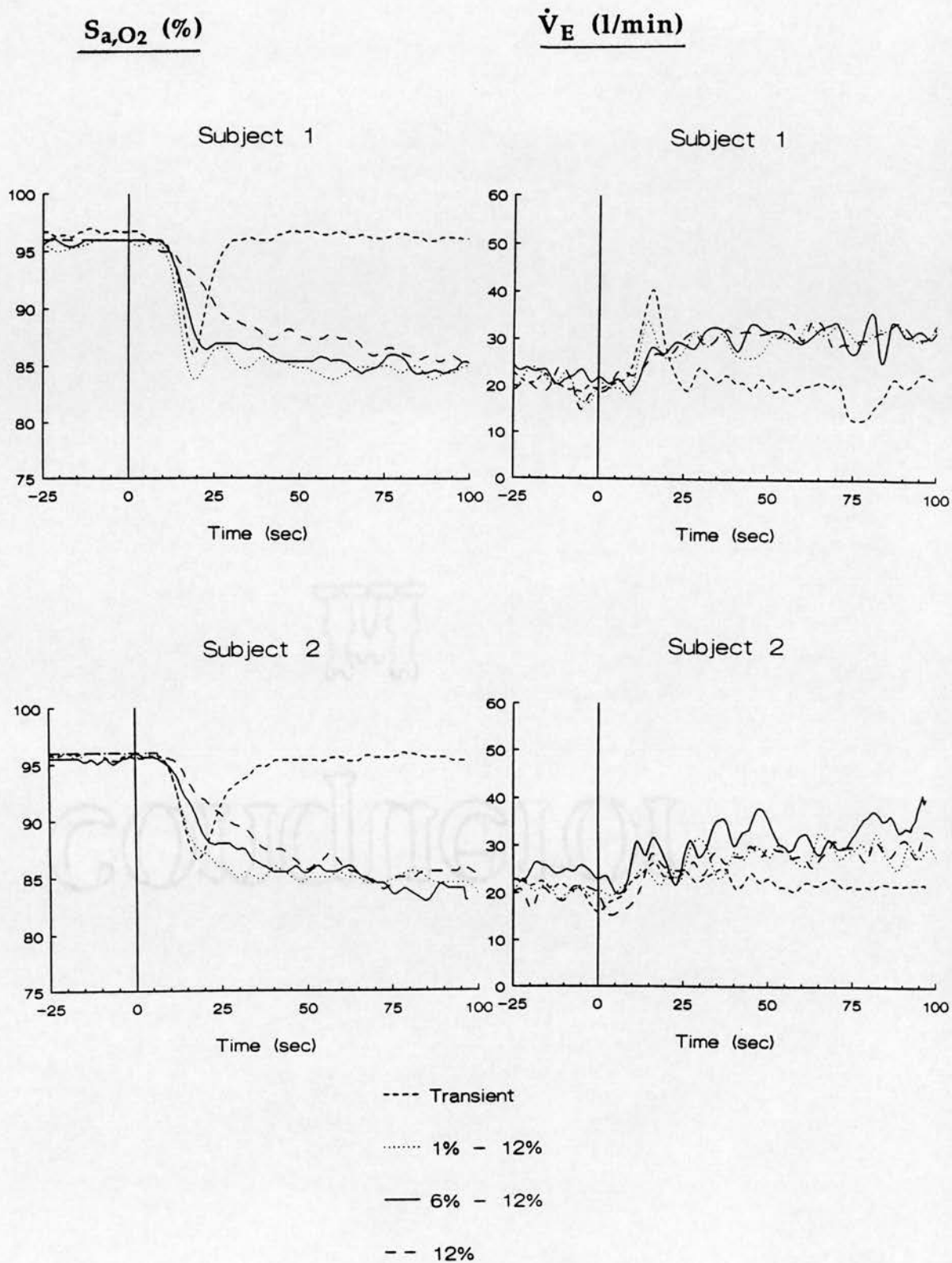


Fig 6.1 Plots of S and V for each of the 5 subjects. Each curve is the average of 2 steps or 6 transients.

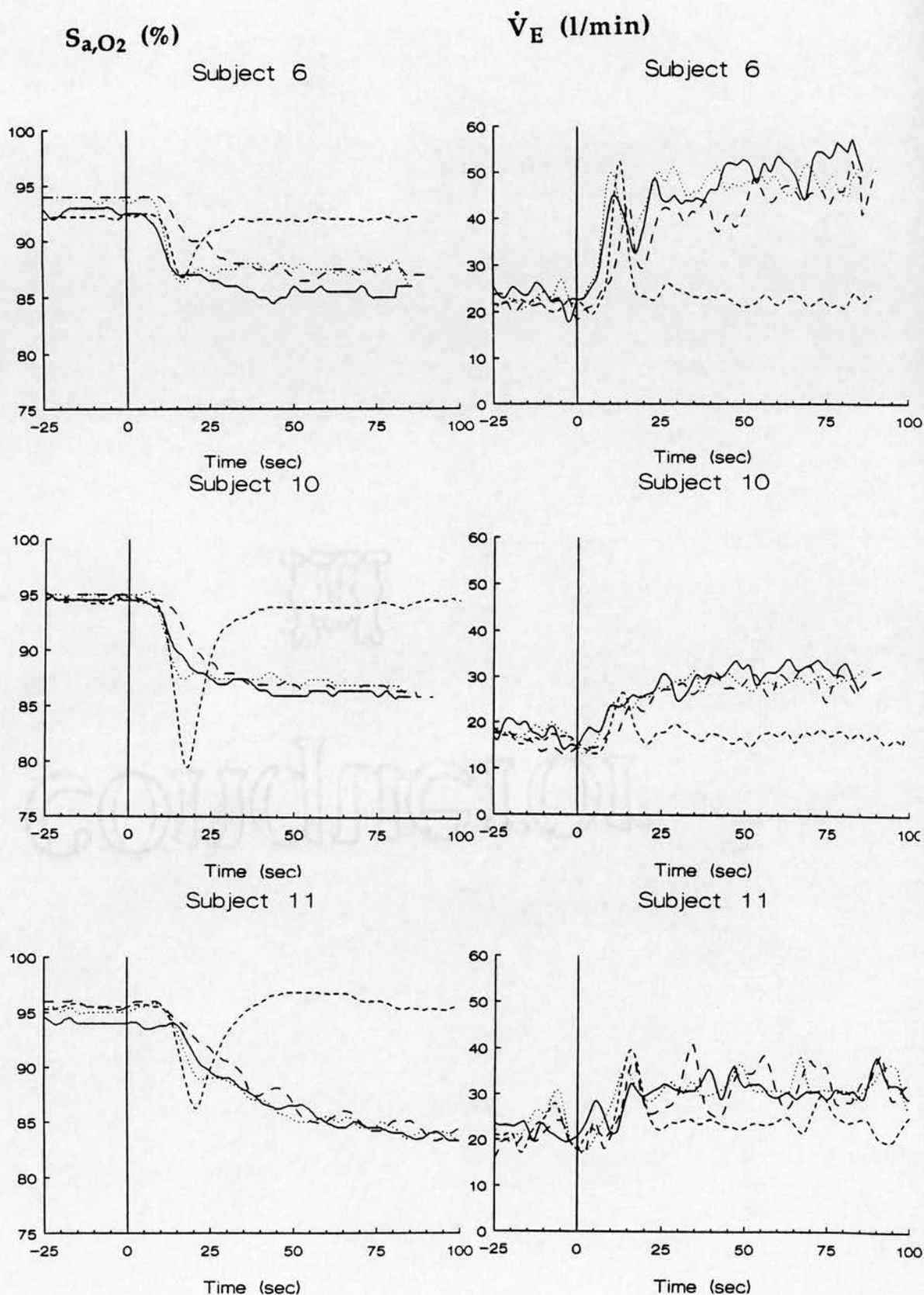
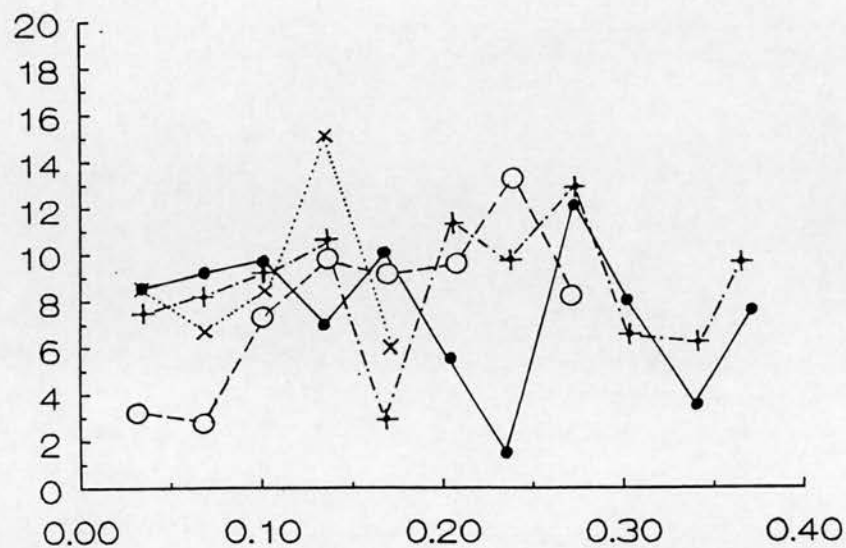


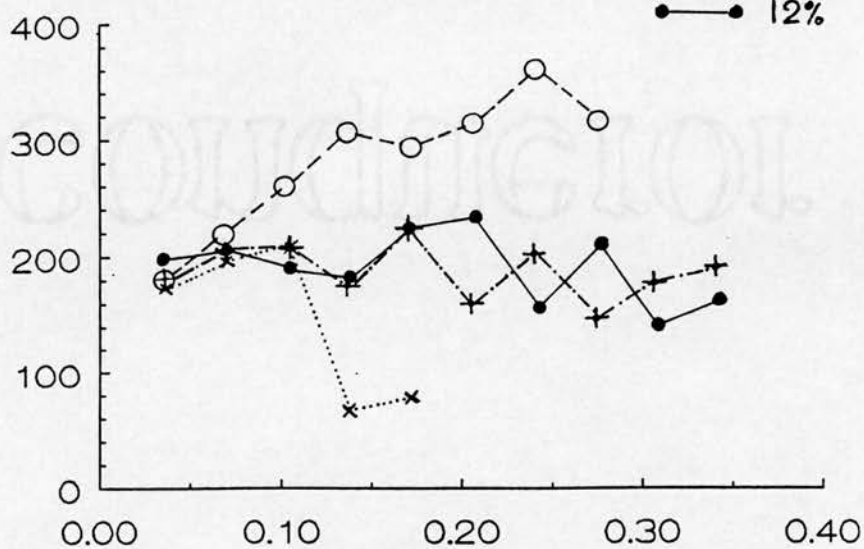
Fig 6.1 cont.

Amplitude (dB)



Frequency (Hz)

Phase Shift (degrees)



Frequency (Hz)

Fig 6.2 *Transfer functions for Subject 10. These were derived from the curves shown in Fig 6.1.*

6.3.1 A test of resolution

The differences found previously between the responses to step and transient hypoxia did not show up clearly as differences between steps with different rates of onset of hypoxia. This could mean either that rate of onset of hypoxia was not an important factor controlling the response, or that the resolution of the experimental method was insufficient to show up the difference.

Two factors might obscure any effect due to rate of change of stimulus: the true rate of fall of S_{a,O_2} is partially hidden by the time constant in response of the ear oximeter; and the inherent randomness in \dot{V}_E causes a low signal-to-noise ratio. To examine these factors, Subject 10 was taken as an example, chosen for her clearly defined response (a high signal-to-noise ratio) and because the difference in her response to steps and transients seen in the earlier study was clear but not extreme.

Firstly, compare the time constant of the oximeter with the rate of fall of S_{a,O_2} . The oximeter's time constant was 1.6 seconds. For Subject 10, S_{a,O_2} fell from 95% to between 86% and 88% by the end of each step. A measure comparable to the "time constant" of response may be made by noting the time taken to fall to about 2/3 of this difference, ie to about 90%. For the interpolated averaged pairs of traces this time was 4, 5, 7 and 12 seconds (for the transients, 1% to 12% steps, 6% to 12% steps and 12% steps respectively). These times, and in particular the difference between the steepest and least steep steps, were large compared with the oximeter time constant, so the distortion due to measurement was unlikely to be important. For this subject, and indeed with all the subjects, the slopes in S_{a,O_2} decreased in the expected order. It is therefore safe to assume that the experimental method was producing true differences in rate of onset of hypoxia in the manner expected, and that the differences were resolvable by the method of measurement used.

The other factor affecting resolution was the "noise" in \dot{V}_E . To examine whether this was large enough to conceal the effect of rate of change, it was first necessary to estimate how large the effect should be. This can be done by extrapolating from the results of the earlier study (described in Chapter

4), the assumption being that the effect remained reasonably constant between the two studies. In the earlier experiment Subject 10's \dot{V}_E rose 10 l/min during transients and 22 l/min during 12% O_2 steps (from a baseline of about 20 l/min) for a fall in S_{a,O_2} of 8%; ie the response to a step was a little over double the response to a transient. The baseline \dot{V}_E for this subject during the current study was somewhat lower, at about 16 l/min, perhaps due to a lower level of exercise. If the difference between the responses to transients and steps in the earlier study were due to rate of change of stimulus, then in this study this subject's response to the 12% O_2 step should be about double the response to the 1% - 12% O_2 step. However, in each case the \dot{V}_E rose by 13 ± 2 l/min. The noise in \dot{V}_E was approximately ± 2 l/min, which was insufficient to conceal the expected difference in response. The experimental method was therefore of high enough resolution to demonstrate the difference had there been one, and so it can be concluded that rate of change of stimulus was not a major factor in determining the overall response of this subject.

The signal-to-noise ratio was lower for the other subjects, but there was no systematic variation in the means of 30 second periods towards the ends of the steps. Again, if rate of change were an important factor in the earlier experiment, then quite marked differences in such mean values would be expected, certainly much greater than the standard errors of these means.

The transient response of Subject 10 in this study was quite different from the step responses. S_{a,O_2} fell approximately 8% during the steps but about 15% during the transients (from a baseline of 95%). \dot{V}_E rose about 13 l/min during the steps and about 11 l/min during the transients, so the rise in \dot{V}_E for a given fall in S_{a,O_2} was more than double for the steps than for the transients, just as in the earlier study.

6.4 DISCUSSION

It is clear from this study that the difference in magnitude of response to transient and step hypoxia was not due primarily to the difference in rate of onset of hypoxia between the two types of stimulus. This is not entirely surprising from the theoretical point of view. Studies (Band et al, 1978; Cross et al, 1984) have suggested that a component of the response of the carotid body to CO_2 depends on the rate of change of P_{a,CO_2} , the response

increasing as rate of change of P_{a,CO_2} increases. It would seem unlikely that the **opposite** should occur for P_{a,O_2} , ie that increasing the rate of change of P_{a,O_2} would **decrease** the response of the carotid body to hypoxia. In most naturally occurring circumstances it would be expected that the rate of change of both P_{a,CO_2} and P_{a,O_2} would change together (for example as exercise level increases or during breath holding): if one were to cause an increase in response of the carotid body and one to cause a decrease, the efficiency of the carotid body in reacting to such changes would be seriously reduced.

A mathematical argument set out in Chapter 8 (§8.6.1) again supports the conclusion. Summarised, this states that it is very hard to conceive of a reasonably simple system where the gain increases as the rate of change decreases, without the system being over-sensitive when the gradient approaches zero (eg towards the minimum level of S_{a,O_2} in a transient).

Two subjects in the first study responded less to steps than to transient hypoxia. This might be due to a rate of change component as their response increased with increasing rate of change, so the theoretical arguments just outlined do not hold. Dutton et al (1973) found such a rate of change component in their experiments on dogs, though for a considerably steeper rate of change (they controlled directly the P_{O_2} of blood perfusing the carotid bodies). If rate of change of stimulus were an important factor in the response of these two subjects, then their response to a 1%-12% O_2 step should show a rapid rise in \dot{V}_E to a peak, before decreasing to the level associated with the 12% O_2 step once the stimulus stopped changing. Unfortunately these subjects were unavailable for testing during the current study

Having ruled out rate of change of S_{a,O_2} as the main factor producing the non-linearity in the \dot{V}_E - S_{a,O_2} system for most subjects, alternative explanations must be sought. Two theoretical approaches can be used for this. The first combines as much relevant current theory as possible and asks whether this can account for the non-linearity. This method is used in the following chapter. The other approach is to experiment mathematically with the relationship found between \dot{V}_E and S_{a,O_2} in an attempt to deduce the nature of the non-linearity, and then to develop

experiments to confirm or refute this. Such an approach is adopted in later chapters.

CHAPTER 7

A PHYSIOLOGICAL COMPUTER MODEL OF RESPIRATION

7.1 INTRODUCTION

The hypoxic steps and transients described in earlier chapters were simulated in a computer model of the physiological processes of respiration developed by Professor KW Saunders (from an earlier model by Grodins). This model contains representations of a large number of the factors affecting the control of ventilation and has been validated against human studies for a variety of experimental conditions. It may be thought of as a concise summary of many of the current quantitative theories of respiration, and so it can be used to see if the difference in response to transients and steps can be explained in terms of existing theory.

If the response calculated by the model for simulations of transient and step hypoxia is similar to the ventilation seen in our human studies, then the differences in response to the two types of stimulus is due to some factor already built into the model but which we have overlooked in simplifying the system to include only S_{a,O_2} and \dot{V}_E as input and output. If the human studies are not well simulated, the non-linearity in the \dot{V}_E - S_{a,O_2} system is due to something not covered in the model, and the search for what that is must continue along alternative lines. Once found, it could itself be built into the model.

7.2 THE MODEL

The model is described in Saunders et al (1980) and Grodins et al (1967) with later changes given in Sarhan et al (1987). These can be referred to for details, and only a broad outline of the concepts and assumptions of the model (as it was in January 1988) will be given here.

The model is mainly concerned with the transport of CO_2 and O_2 round the body in the blood stream. The body is represented as having four basic

sections or "compartments": one for gas exchange (the "Lung"), the others each with their own rate of metabolism (the "Brain", the "Muscle" and "Other Tissue") (Fig 7.1).

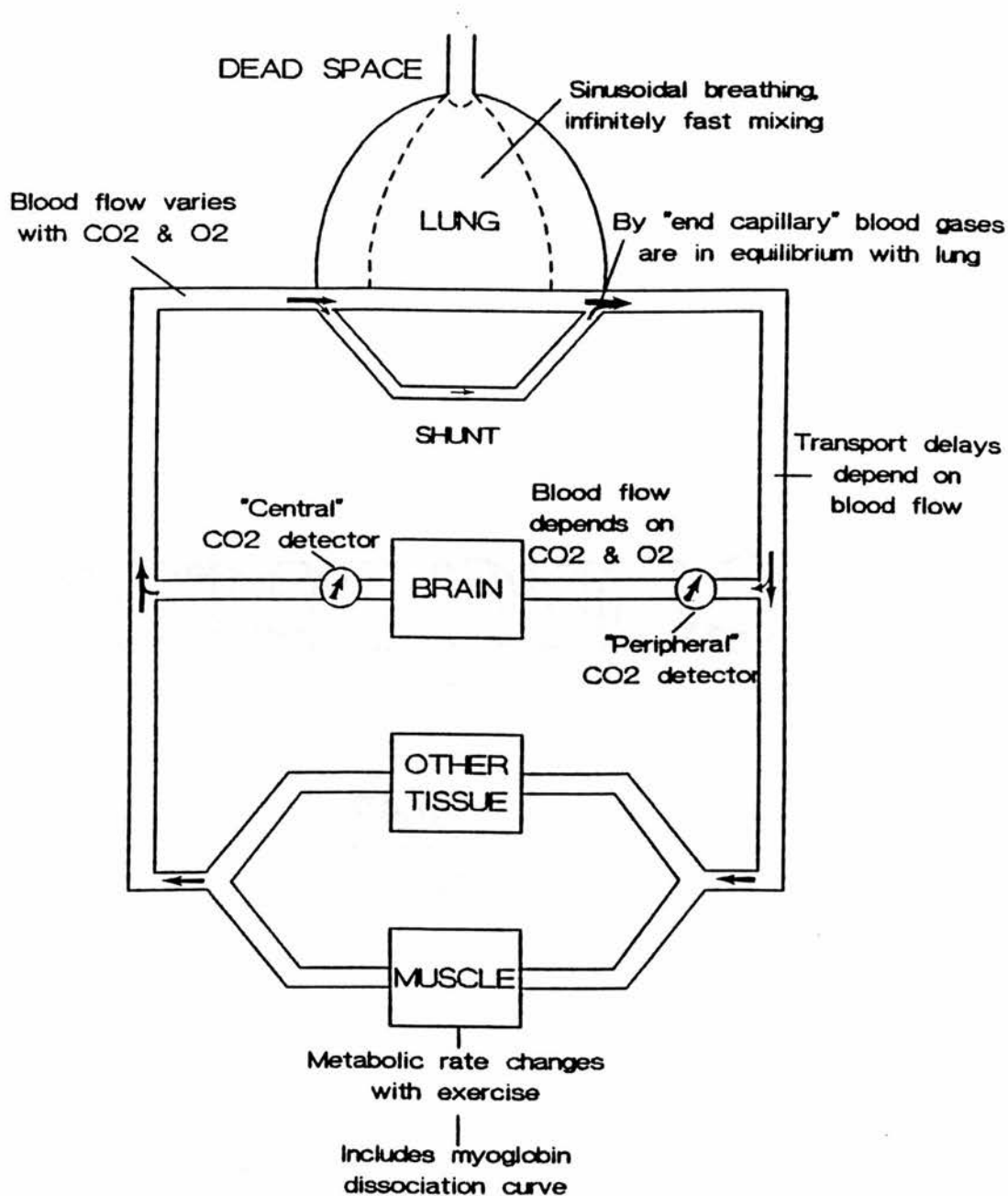


Fig 7.1 Schematic representation of the model used in this chapter.

The Lung is represented as a chamber of varying volume in which infinitely fast mixing of gases occurs. Gas enters and leaves the lung via a small volume of "dead space". It is assumed that the partial pressures of the gases in the blood immediately after passing the lung will be equal to the partial pressures in the lung at that time. 1% of the blood flow bypasses the lung via a shunt, the gas content remaining unchanged.

The Brain has a volume, a metabolic rate and a blood flow dependent on P_{CO_2} , P_{O_2} and cardiac output. In later versions of the model, two detectors of P_{CO_2} , representing the peripheral and central chemoreceptors, are situated at the entrance and exit of the brain compartment: the P_{CO_2} at these points is used to calculate \dot{V}_E .

The Muscle has a volume, a blood flow, and a metabolic rate that is dependent upon the level of exercise. An approximation to the myoglobin dissociation curve has been incorporated in the calculations of \dot{V}_{O_2} .

Other Tissue represents the remainder of the body, the tissue whose metabolism does not change on exercise. This has a volume, a blood flow and a constant rate of metabolism.

The blood carries CO_2 and O_2 round the various compartments. Transport delays are calculated from the moment-by-moment blood flow and volume of vessel travelled through. O_2 is transported both in solution and combined with haemoglobin, the dissociation curves (including Haldane and Bohr effects) being approximated by a complex mathematical equation. The blood flow past the lung and through the shunt, the "cardiac output" is given by an empirical relationship depending on P_{CO_2} and P_{O_2} ; the blood flow in the muscle and other tissue depends on cardiac output and the brain blood flow.

Mass balance equations for O_2 and CO_2 are solved for each of the metabolising compartments, ie the change in concentration of each of these gases in any instant is calculated as the total mass brought in or taken away by the blood plus the mass produced or absorbed in metabolism during that period. For example, the equation for CO_2 for the brain compartment is

$$\frac{\text{rate of change of concentration in brain}}{\text{volume of brain}} = \frac{\text{mass/second created by metabolism} + \text{mass/sec. brought in blood}}{\text{volume of brain}}$$

i.e.

$$d(C_{B,CO_2})/dt = \{ M_{B,CO_2} + \dot{Q}_B \times (C_{a,CO_2}(\text{delayed}) - C_{v,CO_2}) \} / V_B$$

where C is concentration, M is metabolic rate, \dot{Q} is blood flow and the subscripts B, a and v denote Brain, arterial and venous respectively. C_{a,CO_2} is that of the blood leaving the lung delayed by the transport time from lung to brain.

Ventilation is dependent upon the output of the two CO_2 detectors, and is calculated at the start of each breath. Each detector responds linearly to P_{CO_2} with its own time constant and gain (described in more detail below). An empirical formula from Reynolds & Milhorn (1973) is used to find values for **breathing frequency** and **tidal volume**, and the appropriate sinusoidal breathing pattern is generated for that breath.

The mass balance equations, transport delay integrations, concentration to partial pressure conversions and lung volume calculations are solved four times per second of simulated time.

7.2.1 Validation of model:

The validity of the model was checked by comparing the ventilation pattern obtained from simulated experiments, where the inspired fraction of CO_2 entering the lung compartment was changed in various ways, with the pattern obtained from identical experiments on human subjects. These experiments are detailed in the papers referenced above. Good agreement between experiment and model was found with step changes in inspired CO_2 fraction, and the model remained stable for exercise levels up to 100W.

7.2.2 The Ventilation Controller:

The model was originally designed more to examine the workings of the controlled system than the controller, and in consequence the controller equations are not as detailed as they might be. Although earlier versions

of the model had incorporated a hypoxic as well as a CO₂ controller, the current version only detected and acted upon changes in P_{CO₂}.

The ventilation controller in the early version calculated \dot{V}_E from smoothed values of P_{a,O₂} and P_{CO₂} in such a way as to allow a multiplicative interaction between hypercapnia and hypoxia. No attempt was made to build in any time-dependency, the constants of the equations being derived from steady-state experiments (Lloyd et al, 1958). The equations adopted were:

for smoothed P_{CO₂} > 40 mmHg:

$$V = S1 (P_{CO_2} - 38.0)$$

where $S1 = 2.2 \{1.0 + [16.0 / (P_{a,O_2} - 30.0)]\}$

for 20 < P_{CO₂} < 40 mmHg:

$$V = S2 \cdot (P_{CO_2} - 20.0) + 4.0$$

where $S2 = 0.1 (S1 - 2.0)$

for P_{CO₂} < 20 mmHg:

$$V = 4.0$$

Comparisons of model and human \dot{V}_E for hypoxic stimuli using this controller were not as close as for hypercapnic stimuli. In particular, during the on-transient for a step change in inspired O₂ (to 7, 8 and 9% O₂) the model's ventilation increased with half-times 3-5 times slower than those seen in experiments, and gave large overshoots.

In a pilot study, Sarhan (unpublished data) found that simulations of 3 breaths of nitrogen on exercise using this controller led to the model becoming unstable. Sarhan realised that this was due to the P_{a,O₂} dropping to values near or below 30.0 mmHg at which S1 in the equations above has an asymptote. He altered the equations to use S_{a,O₂} rather than P_{a,O₂} by using an empirical formula relating \dot{V}_E and S_{a,O₂}:

$$V = 153 - 1.33 S_{a,O_2}$$

when P_{CO₂} = 46 mmHg. This leads to S1 becoming:

$$S_1 = 19.1 - 0.166 S_{a,O_2}.$$

Comparison of model and experiment for 1, 2 and 3 breaths of nitrogen showed that the model's S_{a,O_2} fell to a much lower value than that found experimentally (eg from 98% to 67% for two breaths of N_2 with a tidal volume of 1.4 l, \dot{V}_E changing from 22 to 53 l/min), but he concluded that this was due to the assumption in the model that mixing of gases in the lung compartment occurs instantaneously.

The controller was later altered to simulate more realistically the detection of and response to CO_2 . P_{CO_2} is detected both peripherally and centrally, and studies by Bellville et al (1979) separated the components of \dot{V}_E associated with each receptor, giving a gain and time constant for each. The model was adapted to incorporate two detectors of P_{CO_2} , one before and one after the brain compartment, each responding linearly to the difference between P_{CO_2} and a reference level but with different time constants, \dot{V}_E being given by the sum of the two outputs. The gains and time constants derived by Bellville were:

Peripheral: gain = 0.72 ± 0.43 l/min/mmHg

time const = 14.8 ± 11.1 s

Central: gain = 1.41 ± 0.61 l/min/mmHg

time const = 180 ± 96 s

Reference $P_{CO_2} = 36.5 \pm 3.8$ mmHg

This version of the model allowed a more detailed examination of dynamic effects as well as being a closer representation of what occurs physiologically. Unfortunately this was achieved at the expense of the hypoxic controller, which was removed from the model. Because the Bellville values were derived from studies at rest, it cannot be assumed that they necessarily are valid on exercise.

It was this last version of the model on which our studies were carried out. It was based on an Amstrad 1512 microcomputer with hard disk.

7.3 USE OF THE MODEL FOR OUR PURPOSES

The aim in using the model was to simulate transient and step hypoxia and see whether the model's ventilatory response to these showed the difference found in human subjects. Using a fully validated model that incorporated so many factors affecting ventilation would have been a short-cut for checking the differences we found against those predicted by the current knowledge of physiology. Unfortunately, the model as it stood was not fully validated for these objectives. There was no tried and tested hypoxic controller, and the CO₂ controller was derived from rest-based experiments.

Before using the model, several adjustments were needed, and so a degree of model validation was inevitably lost. Independent studies were needed to check the new sections of program, but such experiments were not possible (due to time and economic constraints), so the conclusions from this work are necessarily less universal than would be desirable. Instead, further assumptions to those in the model had to be made, and the conclusions drawn are therefore dependent on the validity of these assumptions.

7.3.1 Adjustments to the model:

1. The hypoxic controller: The existing ventilatory controller had two parts representing peripheral and central detection of P_{CO₂}. The main response to hypoxia is mediated via the peripheral chemoreceptor, so the central component of the model's controller was left unaltered while the peripheral component was changed to respond additionally to hypoxia. This approach was preferable to re-adopting the original O₂/CO₂ controller because the latter was derived from steady-state experiments and had no dynamic component. The original controller was never entirely successful for hypoxia and the problem under consideration was, of course, dependent on dynamic effects.

The following assumptions and boundary conditions were used to design the peripheral controller:

- The time constant for the response to hypoxia was assumed to be the same as for the response to changes in P_{CO_2} .
- \dot{V}_E was assumed to be linearly related to S_{a,O_2} . For simplicity, the steady-state form of this relationship was taken to be

$$V = V_0 + H (S_{a,O_2}(0) - S_{a,O_2})$$

where V_0 and $S_{a,O_2}(0)$ were the normoxic steady state values of \dot{V}_E and S_{a,O_2} (at which the model remained in equilibrium) and H , the hypoxic drive, was a constant which could be set as desired.

- Hypoxic and hypercapnic responses interact multiplicatively, as for the earlier version of the controller (see §7.2.2), although the thresholds described there were not re-introduced.
- In normoxic conditions the controller should revert to exactly its previous form.

These points are most simply combined with the original controller by altering the gain of the peripheral component so that it becomes dependent on the fall in S_{a,O_2} . If V_p is the peripheral component of \dot{V}_E , then the controller was given by the equation

$$\tau_p dV_p/dt = -V_p + G_p(P_{CO_2} - \text{Ref.}P_{CO_2})$$

where τ_p is the peripheral time constant and G_p its gain. The gain was changed to the form

$$g = a (b + (100 - S_{a,O_2})).$$

Values for a and b can be obtained to satisfy the boundary conditions, firstly that when $S_{a,O_2} = 98\%$ (the equilibrium normoxic value taken by the model), the value of the gain must be G_p to be compatible with the original controller, so:

$$a (b + (100 - 98)) = G_p ;$$

and secondly that when P_{CO_2} is at its equilibrium value, P_{CO_2eq} , the steady state form of the controller must be of the form

$$V = H (100 - S_{a,O_2}) + c$$

so that:

$$a (P_{CO_2eq} - Ref.P_{CO_2}) = H.$$

Combined, these give values for a and b of

$$a = H / (P_{CO_2eq} - Ref.P_{CO_2})$$

$$b = (G_p / a) - 2.$$

P_{CO_2eq} can be found by sampling P_{CO_2} in the model at a time when no stimulus is being administered and fluctuations following the end of a stimulus (or starting the model up) have ceased. G_p and $Ref.P_{CO_2}$ take the values found by Bellville et al (given in §7.2.2 above), and H can be set as desired (values of 1 and 4 were used as being representative of subjects our studies with a medium and an exaggerated response (respectively), but the actual value was not important given the large intersubject variation). Using these values, V_p in the peripheral part of the controller was replaced by g , and the overall \dot{V}_E was given by the sum of this peripheral component and the component from the central response to CO_2 .

2. Other adjustments: The model as it stood simulated a subject at rest. To change the simulation to exercise, the cardiac output and muscle compartment metabolism had to be increased appropriately. According to the Guide to Normal Values collected by Jones et al (Clinical Exercise Testing, 1975) the average cardiac output for males exercising with a \dot{V}_{O_2} of about 1 l/min should be about 11 l/min (values for \dot{V}_{O_2} 's of 0.8 and 1.2 l/min are given as 9.0 and 12.5 l/min respectively). Using these new values, the model was run without any ventilatory stimulus (ie simulating breathing room air on exercise) until all the values in the model reached equilibrium, when no value showed an upwards or downwards trend over the course of several breaths. The model was then stopped and all the values saved to a file. These were used as the initial values for all future runs to ensure that the model always started from a position of equilibrium appropriate to the exercise level being simulated.

The routines for calculating partial pressures from concentrations and vice versa were written assuming that neither would change rapidly from one calculation to the next. The algorithm used was the Newton-Raphson method (eg Riley, 1974) taking the last value calculated as the first estimate for the current calculation. The very rapid fall in P_{O_2} during a simulated nitrogen transient meant that frequently the change from the last value was so great that using it as a first estimate was inappropriate and caused the calculation to fail (by assigning a negative value to P_{O_2} during the course of an iteration, a state that caused an error to be registered by the program). This fault was corrected by calculating two "look-up tables", one of 20 values over the full range of partial pressures possible with their approximate concentration equivalents, and vice versa for the other. The routines were altered to use the look-up table to obtain a first estimate and then proceed as before with the Newton-Raphson algorithm. Using a look-up table with this number of values proved on average no more time consuming than using the last value each time. This was because frequently fewer iterations were needed by the routine to reach the desired precision due to the better starting estimate. The main cost was in terms of the extra memory required, but this was small relative to the arrays needed by the other parts of the program.

Other small changes were made to the program to improve efficiency. In particular the method of calculating delayed values (eg of the P_{CO_2} reaching the brain compartment), which was particularly time-consuming, was speeded up by a change in algorithm. All such new routines were checked scrupulously to confirm that they produced exactly the same values as their predecessors over the full range of possible circumstances.

A further addition was the development of routines to plot values of S_{a,O_2} and \dot{V}_E on the computer screen during the course of a study. This allowed the user to keep track of how far the program had run, and at the end of a study a hard copy could be taken using a screen dump.

7.3.2 Simulating the control of P_{CO_2} :

To simulate the attempts during human studies to keep end tidal P_{CO_2} constant during hypoxia, a routine was written that increased or decreased

the inspired fraction of CO_2 to the lung compartment according to a simple linear feedback equation. The average $P_{\text{ET},\text{CO}_2}$ over 5 breaths preceding the hypoxic stimulus was taken as a reference value, and the change in inspired CO_2 fraction was a constant multiple of the difference between $P_{\text{ET},\text{CO}_2}$ for the last breath and the reference $P_{\text{ET},\text{CO}_2}$. Various values for the multiplier were tried, but none was entirely successful, with oscillations in \dot{V}_E occurring. These oscillations are not surprising. A delayed feedback loop is in operation: a low $P_{\text{ET},\text{CO}_2}$ will lead to an increase in inspired CO_2 which will lead to a raised P_{CO_2} which will eventually reach the detector and cause an increase in \dot{V}_E and $P_{\text{ET},\text{CO}_2}$, and so after this delay inspired CO_2 will need to be reduced again. In fact, any undamped feedback loop with a delay involved is likely to oscillate. This shows how difficult and risky such techniques are when attempting to keep CO_2 stimuli constant during human studies.

7.3.3 Studies using the model:

The model was used to simulate studies involving three breaths of nitrogen and periods of 15% and 12% inspired oxygen on exercise, starting from a stable baseline as described above. The response to step hypoxia reached equilibrium in less than two minutes, so computing time was saved by lowering the duration of steps from the three minutes used in human experiments to two minutes. The fall in S_{a,O_2} during simulated nitrogen transients (ie 0% O_2) was much greater than that found during human studies, probably because of the assumption in the model of infinitely fast mixing of gases in the Lung compartment (which would reduce the average P_{O_2} of the gas adjacent to the blood when compared with the slower mixing in real lungs). A somewhat more realistic S_{a,O_2} pattern was achieved by using 5% O_2 rather than 0% O_2 in the three breaths. Such a change was permissible as we were more interested in the response to a particular fall in S_{a,O_2} rather than in simulating the mechanics of the experiment.

The resulting \dot{V}_E traces were examined to see if the overall response to steps for a given fall in S_{a,O_2} was greater than the response to transients. If the model gave an accurate representation of a human's response, this difference should be quite clear. The actual values of \dot{V}_E were considered less important than the comparison between the two types of response: it

is simple to adjust parameters to scale the \dot{V}_E to a realistic level, but this would not affect the success or otherwise of the study which rests entirely on whether the difference in responses shows up.

7.4 RESULTS

Examples of the model's S_{a,O_2} trace, P_{ET,CO_2} , and \dot{V}_E response to three breaths of 5% O_2 and two minute periods of 15% and 12% O_2 are shown in Fig 7.2. H , the parameter that gives a measure of hypoxic drive in the controller equation, was set to 4 and 1 respectively.

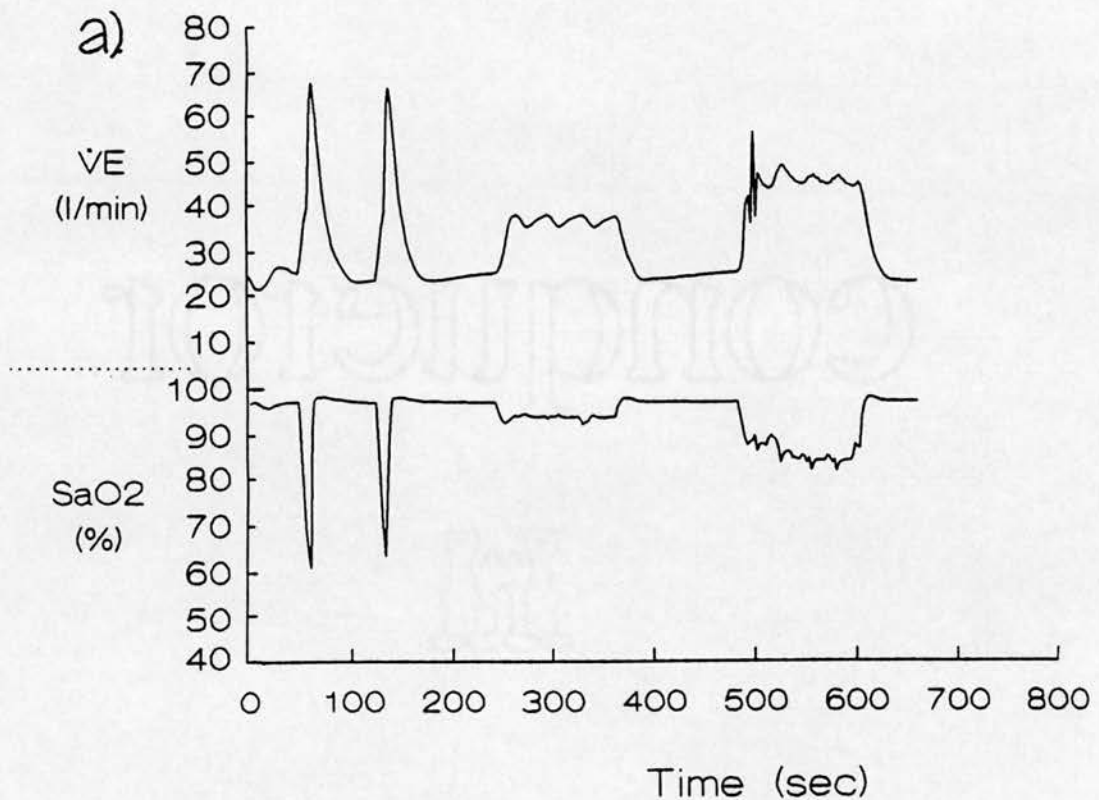
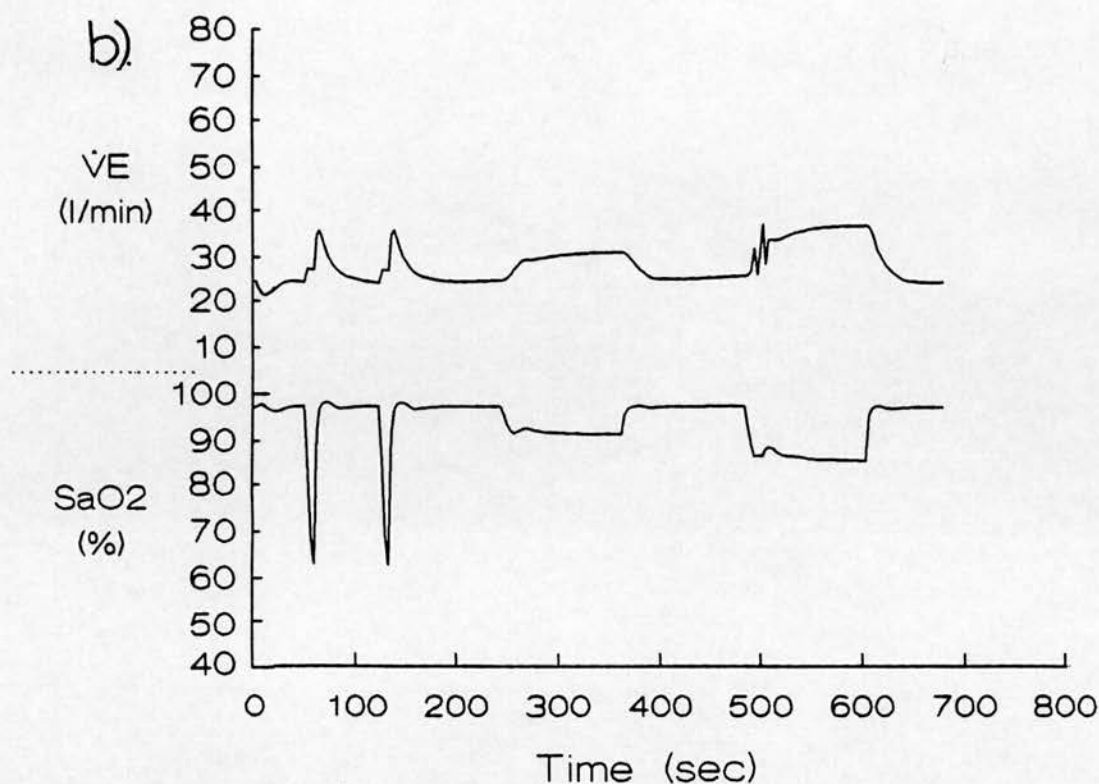


Fig 7.2 \dot{V}_E and S_{a,O_2} output from the model. a) is for $H=4$, b) for $H=1$.



/Fig 7.2 continued

Examination of these traces showed some discrepancies from those obtained during human studies.

1. The S_{a,O_2} fell further during the transients but not as far during the steps compared with a typical human response.
2. The \dot{V}_E during a transient rose continuously over the course of six breaths, and then took approximately twelve breaths to return to baseline. The human response reaches its maximum value in two or three breaths and falls very rapidly at the end of the transient.
3. The response to a step reached equilibrium after about one minute, which was earlier than seen in many human subjects.
4. A short period of rapid oscillation of \dot{V}_E was seen during the first 30 seconds of the 12% O_2 step, though not with the 15% step nor the transients. This was probably due to the attempts at keeping P_{ET,CO_2} constant during the steps: as was described earlier, this is difficult to

accomplish using a feedback system with delays, and the shortcomings will be most obvious during rapid changes.

The difference in response per % fall in S_{a,O_2} at the point of maximum response between a step and a transient was, however, promising. The values when the hypoxic drive parameter in the controller, H , was 4 and 1 were

		Max rise \dot{V}_E (l/min)	Max fall S_{a,O_2} (%)	Ratio (l/min/%)
H=1	Transient	12	35	0.34
	Step	12	15	0.8
H=4	Transient	45	35	1.3
	Step	15	24	1.6

However, the slow rise and fall in the response suggested that the time constant of the controller was unrealistically long, and it could be this that was giving the differences in the ratios. The time constant of the peripheral part of the controller was therefore set at 20 seconds (the central, purely CO_2 controller being left unaltered) and the above values re-calculated for $H=1$ (the case where the difference in ratio was greater).

		Max rise \dot{V}_E (l/min)	Max fall S_{a,O_2} (%)	Ratio (l/min/%)
H=1	Transient	26	34	0.76
	Step	13	12	1.1

The rise and fall in \dot{V}_E during the transients was much more rapid with the faster time constant, and appeared more realistic (Fig 7.3). Unfortunately the difference in ratios was now considerably less than before.

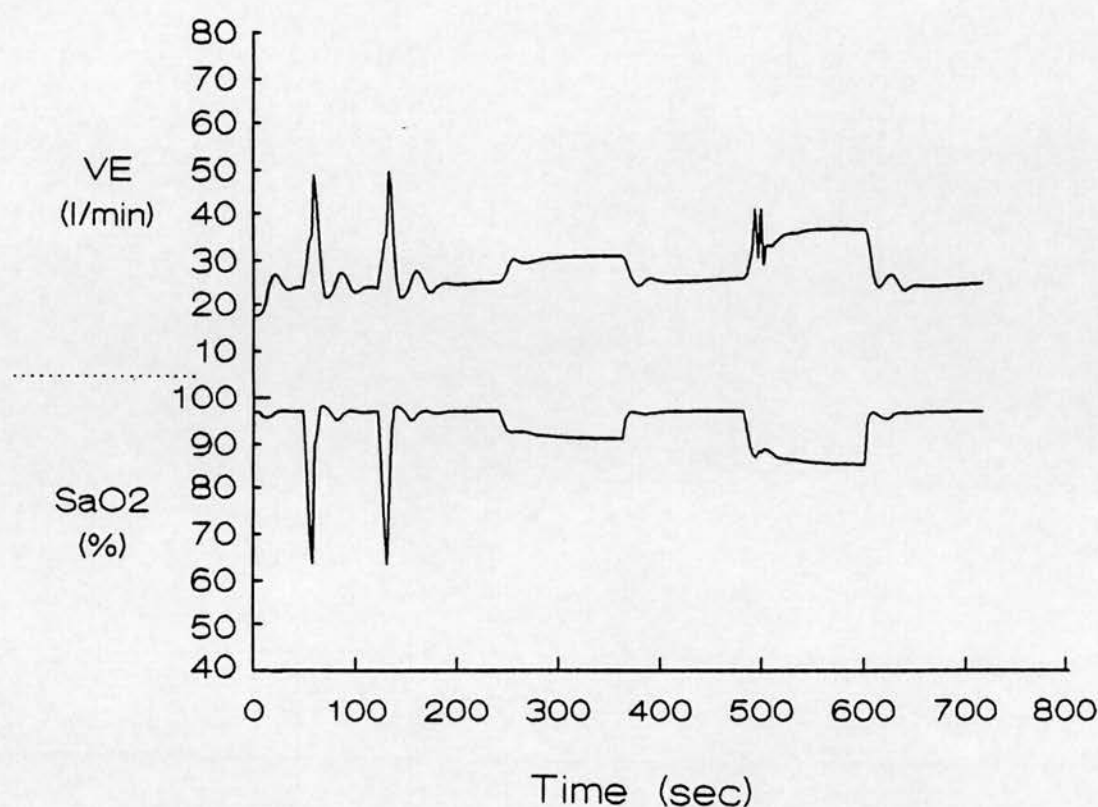


Fig 7.3 *Model output with reduced peripheral detector time constant.*

7.5 DISCUSSION

Even with the faster time constant, a degree of scepticism must be applied to the ratios, especially for the transients. The fall in S_{a,O_2} during the transients was considerably greater than observed in human studies (where S_{a,O_2} typically fell 15%). This is likely to be due to the assumption in the model that gases mix infinitely quickly in the Lung compartment, the slower mixing in real life giving rise to a slower, and therefore shallower, fall in S_{a,O_2} . This is likely to affect the ratio for the transients, increasing it somewhat, as the slowing of the fall in S_{a,O_2} will allow the response to develop more fully (the difference in time constant of response and stimulus will not be so great) leading to a situation where the numerator in the ratio decreases by a smaller amount than the denominator. The difference in ratios between steps and transients will therefore be smaller than obtained from the model.

A further factor to be taken into consideration for the transients is that the oximeter measuring S_{a,O_2} has itself a time constant, and the values of S_{a,O_2} are only given as average values over a breath. Together these will tend to blur the sharpness of the rapid fall in S_{a,O_2} during a transient and will reduce the maximum **measured** fall in S_{a,O_2} relative to the **actual** fall experienced by the subject. The ratios for transients observed in experiments are therefore larger than if compensation had been made for this. Putting this another way, the model, without this blurring to consider, would be expected to show a **greater** difference in the ratios than that seen in experiment.

The small difference in ratios found in the model outputs was probably not just an underestimation of the large difference seen in experiment. It was not possible, however, to reach a definite conclusion on this with the model as it stood. The model parameters might have been those of a subject with a low difference in ratio anyway. This was unlikely given the range of subjects in our experiments for whom a difference was found. More importantly, the discrepancies listed in §7.4 show that the model incorporating the hypoxic controller did not mirror observed responses to hypoxia as accurately as the model with only a CO_2 controller reflected responses to hypercapnia. Less faith could be put in results derived using it until the controller was properly validated. It is quite possible that a better hypoxic controller would have given different responses. Unfortunately, without an obvious physiological basis to build from theoretically, the way to find such a controller would be through experiments such as those we were trying to explain. This would be of no help in the current investigation, of course, as a controller derived from our data would of necessity "explain" our data.

In the chapters that follow a mathematical form for such a controller is derived. If incorporated into the physiological model just described, this might prove useful for future users, and could then be validated against further hypoxic experiments or combined hypoxia/hypercapnia studies.

CHAPTER 8

MATHEMATICAL MODELS BASED ON ONE DIFFERENTIAL EQUATION

8.1 INTRODUCTION

Ideally all manifestations of a subject's hypoxic drive should be expressed by a single formula. This would allow comparison between subjects or between different experiments on a single subject, and it is arguable that our understanding of hypoxic drive is incomplete until this is possible. Applying a simple formula of the form

$$y = mx + b$$

resulted in a different value of m for different types of experiment and so was clearly insufficient (Chapter 4). The Fourier analysis work showed that even a more general linear relationship still could not describe the results of both the step and transient hypoxia simultaneously, so a non-linear element must exist in the relationship. The work with the physiological model (Chapter 7) suggested that this non-linear element might be due to an aspect of the response to hypoxia not contained within that model. Experiments (Chapter 6) failed to show the dependence of the non-linearity on one obvious aspect of the stimulus, its rate of change of onset.

One approach to a clearer understanding of the physiological nature of this non-linearity is to find a mathematical formulation that fits all available data and then to attempt to find physiological mechanisms that could plausibly give rise to such a formula. Experiments could then be devised to test these physiological hypotheses. Frequently there are many possible mathematical formulations that could fit a given data set equally well: wherever a simple relationship holds, it is usually not difficult to postulate a more complicated one that fits just as well to within the accuracy of the data. (As a simple example, points arranged in an approximate straight line might also be considered to lie along a reasonably straight section of a large amplitude sine wave, these two

hypotheses being statistically indistinguishable using the data alone.) It is customary in science to adopt the simpler of two hypotheses that fit data equally well unless there is a strong theoretical reason not to do so. As the physicist Albert Einstein said, a model should be as simple as possible, but no simpler. The aim, therefore, should be to find the simplest formulation relating \dot{V}_E to S_{a,O_2} that fits the data from both transient and step hypoxia experiments.

With no clear quantitative physiological hypothesis to explain the non-linearity, there was no obvious mathematical formula to act as a starting point, but two constraints could be imposed on the form of the relationship:

1. It should approximate to being the linear relationship found by earlier workers (Rebuck and Campbell, 1974) when simulating steady-state experiments.
2. Time must be a term in the equation in order for the dynamic aspects of the relationship, which are clearly important, to be studied.

The starting point chosen for the formulation was therefore the linear differential equation, to which non-linear terms were added to try to improve the fit to the data.

8.2 THE LINEAR DIFFERENTIAL EQUATION

The simplest linear differential equation has the form

$$\tau \cdot dY/dt + Y = gX$$

where Y can be thought of as the response to X , a stimulus varying with time. τ is the time constant of the response, and g is the gain. When the time constant or the rate of change of Y is small, the equation reduces to

$$Y = gX,$$

thus fulfilling constraint 1 above. The response to a sudden change in X is an exponential time-course with time constant τ . A fuller description of the properties of this equation was given in Chapter 1, §1.6.

This simple linear differential equation is widely used in all branches of science (for example in describing chemical reactions or electrical responses). Nothing in nature can respond infinitely fast to a sudden change in stimulus, and frequently the actual change approximates an exponential time-course. It is therefore a natural starting point for our investigation, combining as it does the steady-state proportionality found in the \dot{V}_E - S_{a,O_2} relationship with a plausible dynamic response to a change in stimulus.

8.3 SIMPLE NON-LINEAR TERMS

The Fourier analysis work has already shown that a simple linear differential equation will not be sufficient to describe the data. Rather than make the equation non-linear in an arbitrary mathematical way (eg by squaring X), thereby losing some of the properties described above, the approach adopted was to alter the gain g in ways that mimic types of non-linearity found commonly in nature.

8.3.1 The Saturating Effect:

In real-life systems the output frequently does not remain proportional to the input as this increases, but instead flattens off towards a maximum value. While the stimulus is small the linearity holds approximately, but there is a value of the response which it will never pass. Electronic amplifiers often display this sort of response, as do chemical reactions in many circumstances (one example being the binding of oxygen to haemoglobin in the blood, which increases approximately linearly with rising P_{O_2} in the middle of the range, but flattens off as P_{O_2} reaches high values and the haemoglobin nears full saturation). Because it describes situations where saturation of a substance gives rise to the shape of output, it will be referred to as the "saturating" effect.

A convenient way of describing this shape mathematically is provided by the Michaelis-Menten equation, developed for the field of enzyme mechanics (Carson et al, 1983):

$$Y = aX / (X + b) .$$

This is economical in its use of parameters, requiring only two: a and b . While X is low b will dominate the denominator and the equation will simplify to

$$Y \approx (a/b) X,$$

ie Y is approximately proportional to X with a gain of a/b . As X increases, it exerts more influence on the denominator, and the gain will be reduced. When X is very much larger than b , b will have little effect and the equation will reduce to

$$Y \approx a$$

has reached its maximum value of a and further increases in X will not increase Y further.

This formula can be rearranged into a more convenient form by dividing top and bottom by b , then substituting

$$g = a/b$$

$$\sigma = 1/b$$

to get

$$Y = gX / (1 + \sigma X).$$

Now σ is a direct measure of the degree of saturation (when σ is 0 the equation reduces to a linear one, and as σ increases, so does the distortion). If X or σ is small, g acts like the gain in a linear relationship. The maximum possible value is now g/σ . Plots of this type of equation with different values of g and σ are shown in Fig 8.1.

This formulation can be incorporated into the linear differential equation described above by replacing the right hand side (gX) with the right hand side of the adapted Michaelis-Menten equation:

$$\tau \cdot dY/dt + Y = gX / (1 + \sigma X).$$

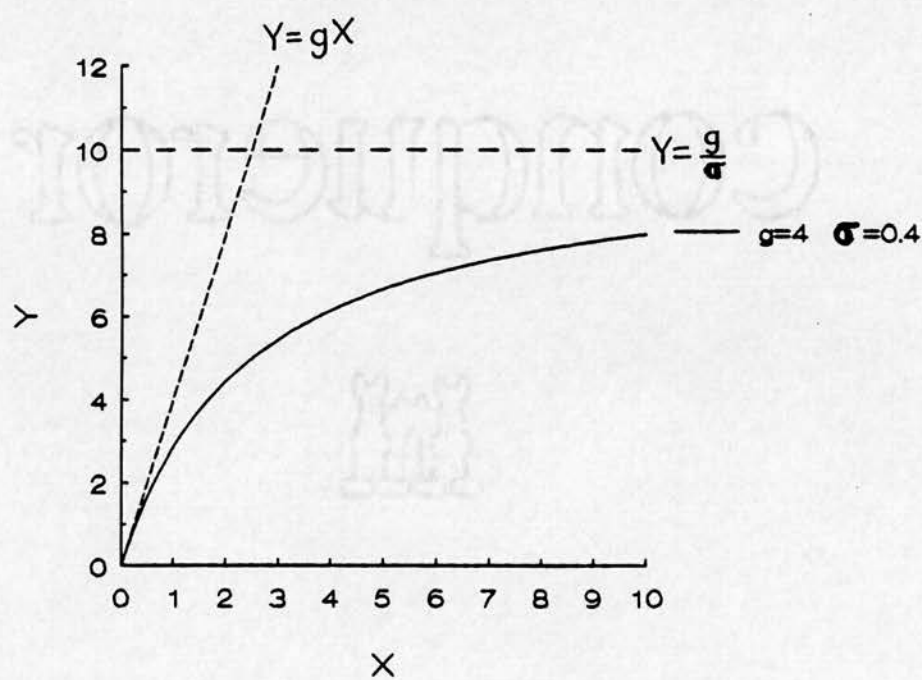
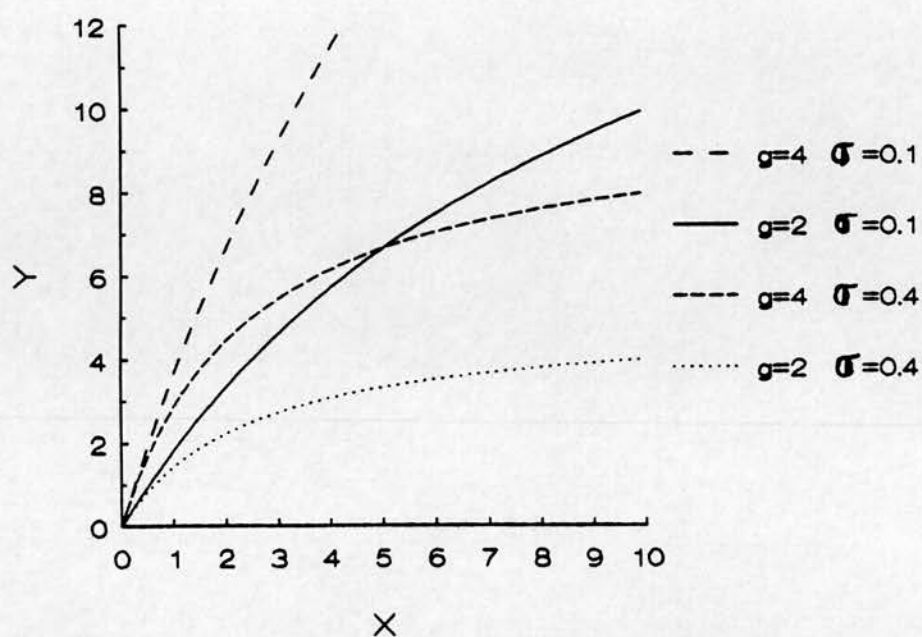


Fig 8.1 The "Saturating Effect" Plots of Y vs X for different values of g and s . The second plot shows how the curve relates to the values of the parameters.

8.3.2 Threshold effects:

The saturating effect is one where the gain varies smoothly with increasing X . Often, changes in the gain of a relationship happen suddenly at a particular level of input. Common examples of this are numerous, including the firing rate of nerve clusters which increases suddenly when the total activity of incoming signals reaches a threshold level. The recently developed branch of mathematics of Catastrophe Theory has identified threshold situations occurring in all aspects of science and sociology.

Threshold effects may be easily incorporated into the differential equation. It is necessary to specify the values of the gain below and above the threshold, and also the threshold itself at which the change takes place. Three parameters are therefore necessary, eg g_1 , g_2 and T in (Fig 8.2):

$$\tau \cdot dY/dt + Y = g_1 \cdot X \quad X < T$$

$$\tau \cdot dY/dt + Y = g_2 \cdot X \quad X \geq T.$$

It is of course possible for the change in gain to occur at a threshold value in a variable other than the input X , eg the rate of change of X (a search for such a threshold was described in Chapter 6).

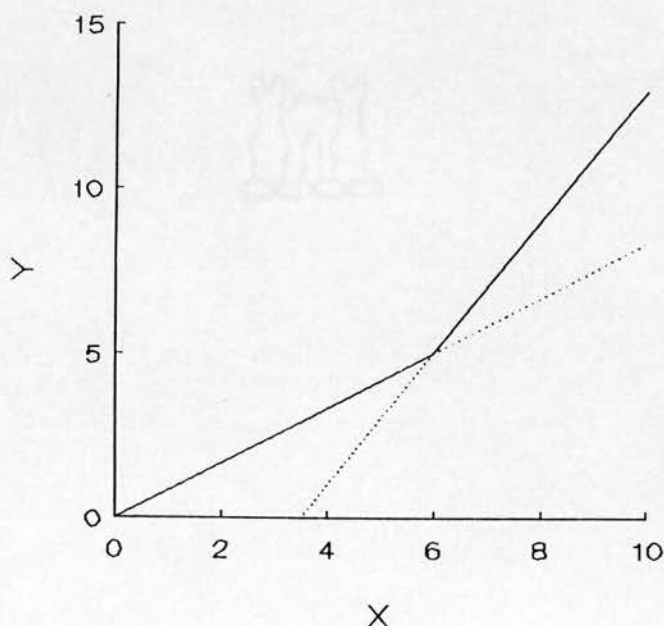


Fig 8.2 "Threshold Effect" The plot shows how the gain increases at a threshold of $X=6$

8.3.3 Other non-linear effects:

Saturating and threshold effects cover a wide range of plausible outputs for naturally occurring systems. It is unlikely that the gain of a system will increase indefinitely with increases in X (Fig 8.3) as high values of X will eventually lead to unrealistic increases in sensitivity. (Such an upward curve is, however, possible for a limited period and is frequently followed by the saturating effect for higher X : this gives the S-shaped response seen in relationships such as between P_{a,O_2} and S_{a,O_2} . The initial part of the curve may be thought of as a saturating effect in reverse and can be expressed mathematically using the same formulation as before but multiplying X in the denominator by -1 (and not allowing its magnitude to reach σ). More complicated curves than this are not likely in a simple stimulus-response system unless additional factors are affecting the response: analysis is then better directed at discovering those factors rather than designing more complex formulae. One example of this, which will be returned to in Chapter 9, is inhibition or potentiation of the gain by something occurring outwith the immediate system being described by the differential equation.

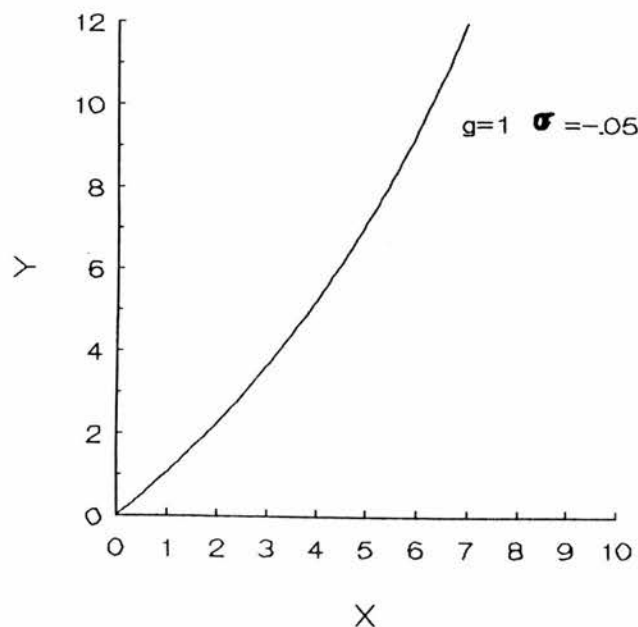


Fig 8.3 Output for constantly increasing gain

Delays: Delays within a system can be a source of non-linearity; and here the non-linearity does not depend upon the alteration of the gain of the equation. If X is delayed by a fixed time before being acted upon, then plotting Y against the undelayed X will prove to be non-linear (Fig 8.4 shows a simple example). Because of this, wherever a delay in the system is known about it must be corrected for before further analysis; but there always remains the possibility that there are delays in the system not known about but which could be affecting the shape of the response.

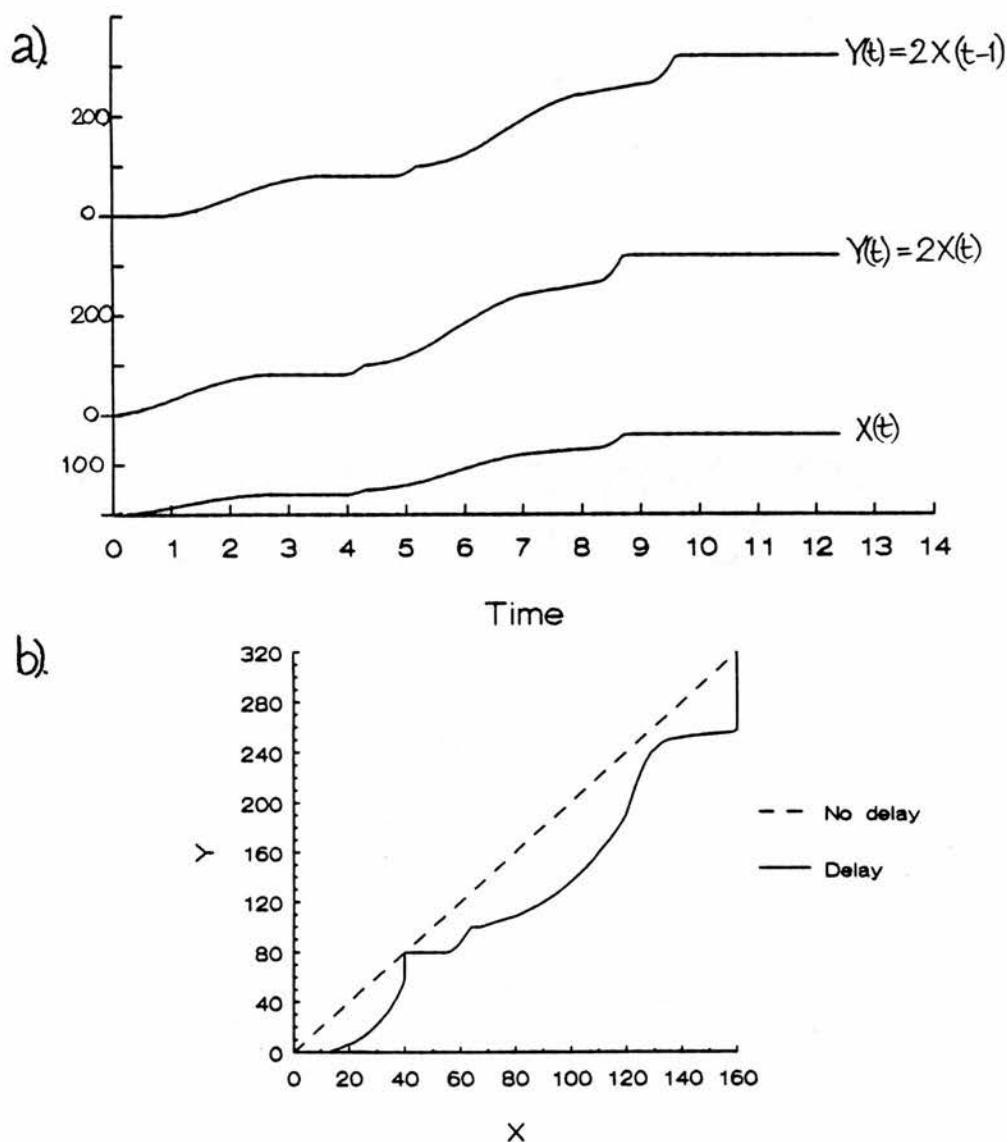


Fig 8.4 Non-linearity due to a delay. *a)* shows a stimulus X and two responses, both equal to $2X$ but the upper one is delayed by 1 sec. *b)* shows the X - Y plot for the delayed response (solid line) and for the non-delayed response (broken line).

8.4 METHOD

The linear differential equation, and differential equations containing the non-linear terms described above, were set up for the relationship between:

X - the fall in S_{a,O_2} from its steady-state normoxic value,

and Y - the increase of \dot{V}_E above its normoxic value.

It was necessary to take the differences from the normoxic values for S_{a,O_2} and \dot{V}_E rather than their absolute values because the differential equation in the form given is only valid where both stimulus and response have the same sign (eg both positive), and the solutions are simplified by having the response zero for zero stimulus.

These equations were solved for a given set of values of S_{a,O_2} (using a simple computer program) to give a set of calculated values of \dot{V}_E (\dot{V}_{Ecalc}). These were then compared with the actual \dot{V}_E (\dot{V}_{Eobs}) for that same S_{a,O_2} trace. The aim was to find a form of the equation with non-linear terms that produced a \dot{V}_{Ecalc} trace similar in shape and size to \dot{V}_{Eobs} . If such an equation were discovered, it would be possible to calculate the parameter values (ie values for τ , g , σ etc.) applying to each subject.

Another approach used was to take an idealised S_{a,O_2} trace consisting of one transient fall and one step fall in S_{a,O_2} , both to the same minimum value (Fig 8.5a). Using this as the input showed clearly what each model was doing, and the task was simplified to a search for a model which resulted in double the rise in calculated \dot{V}_E during the step compared with during the transient. Without the constraint of having to fit real data, it was no longer necessary to worry about obtaining particular values for \dot{V}_E , as once the relative difference between transient and step response had been obtained, the response could be scaled as necessary by changes in the gain of the equation. The process was thereby speeded up considerably. If an appropriate relative difference were found, the resulting model could be tested on real data.

Each of the following non-linear effects were tried:

1. Saturating effect
2. Change in gain when input crossed a threshold
3. Change in gain when output crossed a threshold
4. Change in gain when $d(S_{a,O_2})/dt$ crossed a threshold
5. Change in gain when $d(\dot{V}_E)/dt$ crossed a threshold
6. Different gains for positive and negative slopes of input
7. Gain changing at different thresholds for positive and negative slopes of input
8. Change in time constant when input crossed a threshold
9. Change in time constant as for 2. to 7. above

The next chapter describes a much more powerful and satisfactory program than the one used for this work. That program was used to check the results of this chapter's investigation, and it produced superior plots which will be shown in the next chapter. Because of this, and because there was considerable overlap between the two programs, discussion of the details of the program and its outputs will be left to the next chapter, and the discussion here will be based mainly upon a theoretical approach developed after much time had been spent with the simpler program without any consistent success at simulating the difference in response to steps and transients.

8.5 RESULTS

Fig 8.5 shows examples of some of the model outputs for the idealised S_{a,O_2} input. It was possible to obtain double the increase in the response to the step compared with the response to the transient if the time constant was greater than 12 sec, but this resulted in a very slow return to baseline following the stimulus and a slow initial response at the onset of stimulus (Fig 8.5d&e). Using non-linear terms did not improve this, with the exception of one rather contrived method where a threshold change in the gain was set so that the gain increased once the response reached a certain

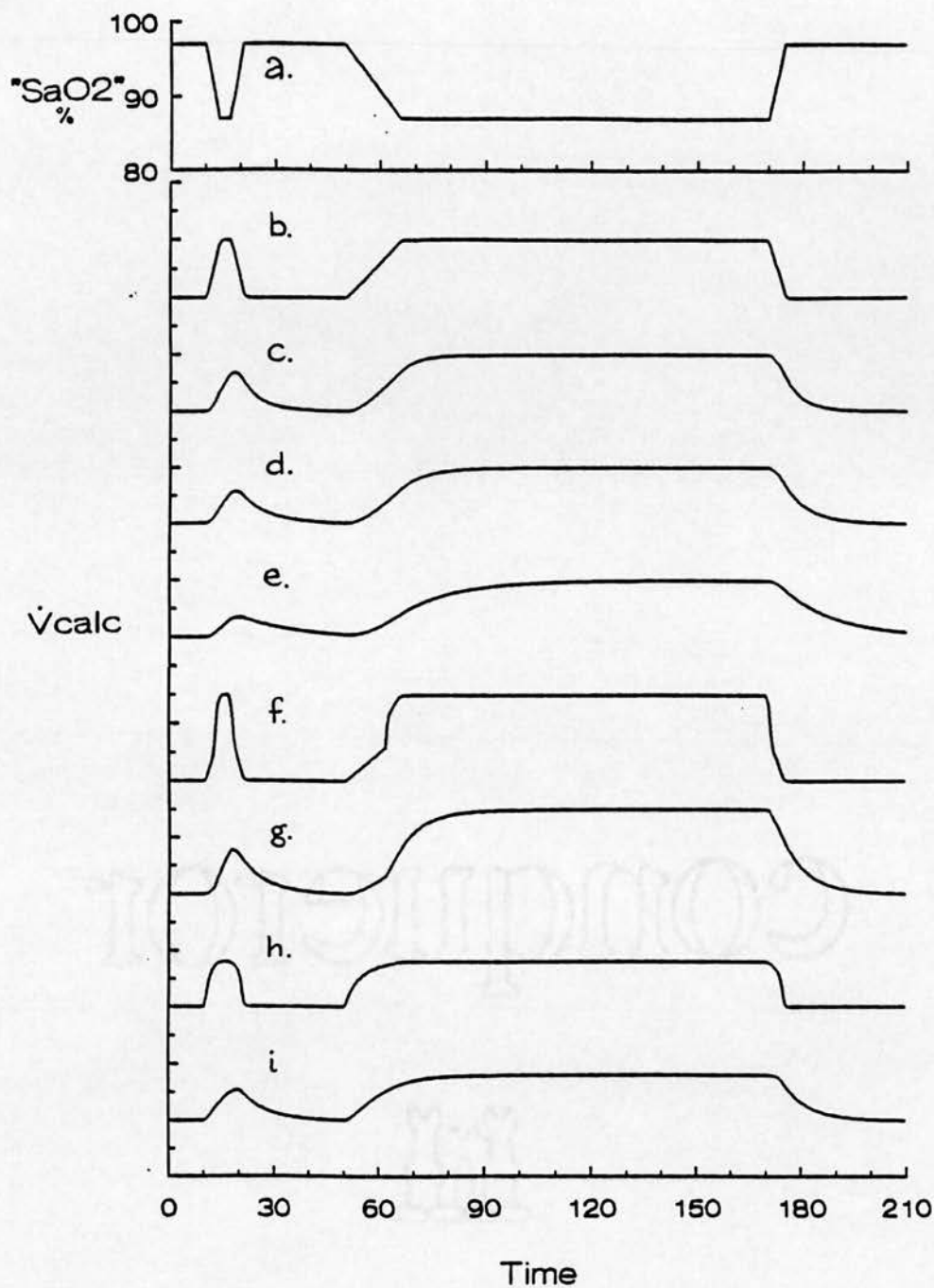


Fig 8.5 Responses of the simple model to an idealised S_{a,O_2} input. *a* shows the stimulus; *b, c, d* & *e* show the response of a linear model with time constants of 1, 5, 10 and 15 seconds respectively; *f* and *g* are from a model incorporating a threshold change in gain with time constants of 5 and 10 seconds; *h* and *i* are from a model incorporating the saturating effect, again with time constants of 5 and 10 seconds.

level, then adjusting the time constant so that the transient response never reached this threshold level, whilst the step response did. This resulted in a greater response to steps without requiring as long a time constant as in the simple linear model, thereby reducing the tail-off following the stimulus. Unfortunately this model was ineffective with real data, as occasionally the response to a transient would exceed the threshold simply because the fall in S_{a,O_2} was greater than that for the steps: the differentiation between step and transient response would then vanish. In general the non-linear effects did not help with the difference between step and transient response, and often caused unrealistic distortion in the response (Fig 8.5f,g,h,i).

8.6 DISCUSSION

The inability of all of the differential equations to reflect the observed difference between response to steps and transients can be explained by considering the problem from the opposite direction. What form would the equation have to take to produce a much greater output for a given fall in S_{a,O_2} during step changes than during transients?

If the time constant is short, as it must be to give a sharp rise in \dot{V}_E for the transient, a linear equation will produce an output similar in shape to the input trace, though reflected and magnified by the value of the gain (Fig 8.5a). In particular, the maximum outputs reached during the transient and during the step would be the same.

The non-linearity in the equation must cause the gain to be double (on average) during the step compared with during the transient. This is not possible for the types of non-linear terms where the gain depends upon the value of the input as the input passes through the same range of values in each case, and so both step and transient will be affected by the change in gain to the same degree. Figs 8.5h&i show an example where the gain displays a saturating effect: the maximum value attained during each type of stimulus is still identical. Because the output so closely mirrors the input when the time constant is small, the same will be the case where the gain changes according to the level of output.

This shows that it is necessary for the non-linearity to depend upon something that differs between the two types of stimuli. In the example there are two differences between the step and the transient: their durations, and the rate of change of S_{a,O_2} in the initial part of the stimulus.

8.6.1 Rate of change:

The experimental work of Chapter 6 suggested that the rate of change of S_{a,O_2} is not an important factor deciding magnitude of response. Some examples here will back this up, showing that it is mathematically very difficult to construct an appropriate controller equation varying with slope.

The slope of the onset of the transient stimulus is steeper than that of the step. If the non-linearity is to be based upon this, the gain of the equation will have to be greater for **lower** rates of change, and, in the example, the gain must on average be double during the low rate of change step than during the transient. For simplicity, assume that the average slope during the onset of the transient is double that for the step. This would suggest either that the gain is roughly inversely proportional to the slope, or that there is a threshold slope (somewhere between the transient and step slopes) below which the gain suddenly doubles. Both these ideas seem plausible.

The respiratory system cannot, however, be reacting to average slope, as this requires the entire time-course to be known before the response starts, and yet the response clearly starts almost immediately the stimulus is detected. The response would have to be reacting to moment-by-moment (or possibly breath-by-breath) slopes. This causes problems with both the suggested slope-gain relationships. There has to be a period as the stimulus reaches its maximum level where the rate of change levels out, and eventually becomes zero.

If gain were inversely proportional to slope, as suggested, then the gain would be very high indeed in this area of low rate of change, and the response would be seen to increase dramatically in magnitude as the stimulus reached its peak. Clearly there must be a maximum possible gain

for low and zero slopes. At the maximum level of stimulus, both the transient and the step will have a slope so low that the maximum gain has been reached. The gain for the two types of stimulus will therefore be the same, and unless the time constant is quite long, the response to the transient will be seen to increase at the end towards the level reached by the step, cancelling out the effects of the lower gain earlier in the response. To keep the difference in magnitudes at the end, the time constant must therefore be long compared with the duration of the low and zero slope section of the transient stimulus. Such a slow time constant conflicts with the fast initial response to the stimulus seen in the studies.

A similar argument applies to the suggestion that there might be a threshold slope above which the gain halves. Towards the end of the stimulus the transient's slope must cross back below the threshold, and at that point there should be a sharp increase in response. This is not seen.

It might be possible to find just the time constant necessary to mask the changes in gain at the end of the onset of the transient stimulus, and then to construct a curve relating gain to slope that starts at a maximum and then curves down steeply enough to end up with the overall gain for transients being half that for steps given the durations of stimulus and the value of the time constant, but the balancing involved in such a construction is unlikely to be effective if the shape of stimulus alters even slightly. The effect to be explained is not constrained to ideal shapes of stimulus, but occurs over a range of similar but not identical shapes. Basing a model around rate of change of stimulus therefore seems to be too complicated.

8.6.2 Duration:

The other difference between the step and the transient is the duration of the hypoxia. The simplest way to exploit this difference using the linear differential equation is to make the time constant large compared with the duration of a transient. During the transient there is then insufficient time for the response to achieve more than a fraction of its maximum level before the hypoxia ends, whereas during a step there is sufficient time for the maximum level of response to be reached. If the time constant is made large enough, it is possible for the maximum response to the step in the example being considered to be twice that to the transient (Fig 8.5d&e).

However, the Fourier work (Chapter 5) showed that the system was not linear, so there must be some non-linearity in the equation in addition to the large time constant. Fig 8.5e shows the output of a linear system with a large time constant. The shape of the output does not match that of the observed \dot{V}_E : in particular, the rise in \dot{V}_E at the start of a transient is much slower in the model's output, and there is a long "tail" following each period of hypoxia while the response returns exponentially to baseline. A large time constant may be part of the equation, but it is not the whole answer.

There are other ways in which the differential equation might be modified to capitalise on the difference in duration between the two types of stimulus. If the gain is allowed to increase with time following the start of hypoxia, the result will be a larger response to steps than transients. The following equation would continue to increase indefinitely during a step:

$$\tau dY/dt + Y = (a + (t-t_0)b) X$$

where the gain has been replaced by a quantity that increases linearly with time after the beginning of hypoxia (t_0). However, observed responses frequently flatten off after a few minutes of hypoxia, so the expression giving the gain as a function of time must do likewise. The equation must therefore be quite complex: one possibility might be

$$\tau dY/dt + Y = g(t) X$$

where

$$g(t) = b + (b-a) \exp(-t/\tau_2) ,$$

ie the gain starts at a value $g = a$ then increases exponentially towards a maximum value $g = b$ with a time constant τ_2 . By manipulating values of a , b , τ and τ_2 , such an equation may be made to fit the sort of shape found in the experimental data. If τ is small and τ_2 in the order of a few minutes, then the response to transients will be approximately the fall in S_{a,O_2} times a , and the final response to steps will be the fall in S_{a,O_2} times b . This would seem to be a solution to the problem, but there are problems with relating this equation to physiological mechanisms. Some factor is required to start increasing when the S_{a,O_2} falls below its normoxic value and then for the increase gradually to slow down until it reaches a final

value, the time-course being purely dependent upon time. The process starts as soon as hypoxia is detected, but then runs on independent of degree of hypoxia or anything else. This seems unlikely. It would be easier to imagine this factor responding to some change that was occurring in the body rather than simply changing with the time from a difficult to establish point (the moment that hypoxia started). As the main changes occurring are the fall in oxygen level in the blood and the ventilatory response to this, it might be possible to develop an equation that is based on these changes and works along the same lines as the equation given above. The exploration of this possibility is the subject of the next chapter.

8.7 SUMMARY:

A simple linear differential equation was chosen as a natural starting point for a mathematical investigation of the hypoxic response to transients and steps. A variety of non-linear terms commonly found in naturally occurring systems were incorporated into the equation in an attempt to find an equation that gave a greater response to steps than transients, without success.

It was argued that the equation must be highly dependent upon either the rate of change of stimulus or the duration of stimulus in order to produce such differences. Theory and experiment suggest that rate of change is not important, so the differences must be due to the duration of hypoxia. This is most simply built into the equation using a large time constant, but this is not the whole solution. An equation with two time constants provides a better solution.

Rather than make a single differential equation progressively more complex, the next step was to develop a solution based on two linear differential equations.

CHAPTER 9

THE DEVELOPMENT OF A COMPUTER PROGRAM FOR INVESTIGATING A SYSTEM BASED ON TWO DIFFERENTIAL EQUATIONS

9.1 INTRODUCTION

The search for the simplest mathematical description of hypoxic drive compatible with our data required more than a simple non-linear differential equation. The conclusion of the last chapter was that a search progressing logically through more and more complex descriptions should turn next to an investigation of a system based on two differential equations. As will be seen, two differential equations can be combined in a variety of ways, and the output for such a system is, in all but the simplest cases, too complex to be analysed without a computer. This chapter describes the development of a large computer program written to solve such equations, and then outlines the way such a program should be used. The next chapter will describe the actual use this program was put to, and the model that was found to fit the data.

This chapter is necessarily long and technical, and considerable detail is given to allow proper critical assessment of the following chapters. The program was developed over a long period of time and is perhaps the essence of the thesis, containing both established mathematical techniques and original solutions to problems.

9.2 TWO DIFFERENTIAL EQUATIONS

The differential equations are of the same type as described in the last chapter. The linear form of each is defined by two parameters, the gain and the time constant, and these can, of course, have different values in the two equations. Delays and a Saturating Effect of each gain were allowed for in each equation, though were not always used.

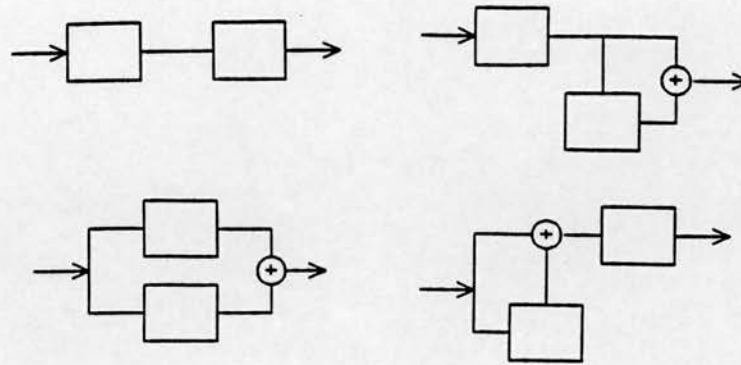


Fig 9.1 Various arrangements of two equations (each represented by a box with an input and an output).

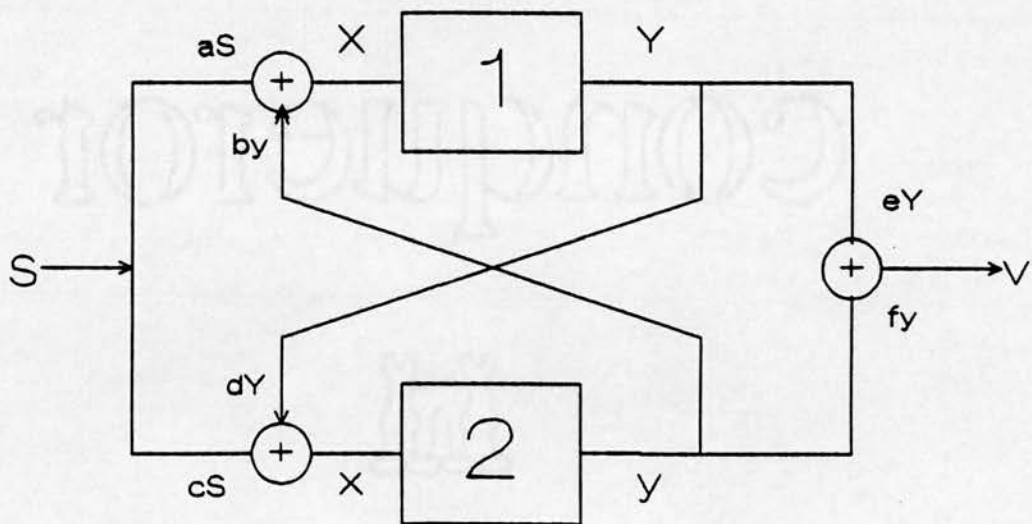


Fig 9.2 The generalised arrangement of equations (see text below).

If each equation is thought of as a box with an input and output, it may readily be seen that several arrangements of two such boxes are possible (Fig 9.1). The boxes may be arranged sequentially, the output of the first being the input to the second; or they may be in parallel, each with its own input and output, the two outputs being combined in some way (eg added

or multiplied) to get the output of the system as a whole. There is a whole range of possibilities between these two extremes, where the input of one or both boxes is a combination of the input to the system and the output of the other box. The completely general case is shown in Fig 9.2. If the input and output of one box are X and Y respectively, and of the other x and y, and the input to the system is a value S, then:

$$X = (aS + by) / (a+b)$$

$$x = (cS + dY) / (c+d)$$

and the overall output V would be given by

$$V = (eY + fy) / (e+f)$$

(or $V = Yy$ if the outputs are to be multiplied). The division by the sum of the parameters in each equation makes sure that only the gains of the differential equations affect the magnitude of response. The arrangement parameters a to f can be manipulated to give any desired arrangement of the boxes. For example, if b, c and e are all zero then the arrangement is the sequential one; and if b and d are zero the arrangement is the simplest parallel form.

9.2.1 Inhibition/Potentiation:

A further form of simple non-linearity can be introduced into a system of two differential equations. The gain of one equation can be increased or decreased according to the level of output of the other, which corresponds to one process "**potentiating**" or "**inhibiting**" the other. This sort of effect can often be seen in neural systems. There are many ways in which potentiation or inhibition can be expressed mathematically. One of the simplest ways is to multiply the gain in the linear differential equation by $(1-IY)$, so that

$$\tau \cdot dy/dt + y = (1-IY)gx$$

where Y is the output of the other equation and I is a parameter measuring inhibition. Generally I will be small so that IY will be much less than 1: the greater the value of I, the greater the inhibition. If I is

negative, then potentiation will occur, ie the gain will be increased as the output of the second equation increases.

9.2.2 Definition of the generalised model:

The following set of equations summarises all of the above, and is the generalised model written into the program.

$$\dot{V}_E(\text{model}) = V + \dot{V}_E(\text{normoxia})$$

$$V = (eY + fy) / (e+f) \quad [\text{or } V = Yy]$$

$$\tau_1.dY/dt + Y = G_1.X$$

$$\tau_2.dy/dt + y = G_2.x$$

$$X = (aS + by) / (a+b)$$

$$x = (cS + dY) / (c+d)$$

$$G_1 = (1-I_y)G / (1 + \sigma_1.X)$$

$$G_2 = (1-iY)g / (1 + \sigma_2.x)$$

$$S = S_{a,O_2}(\text{normoxia}) - S_{a,O_2}$$

Key: $\dot{V}_E(\text{model})$ - the model's calculated ventilation

S_{a,O_2} - a subject's breath-by-breath S_{a,O_2} trace

(normoxia) - the subject's average value pre-stimulus

S - the stimulus, ie the fall in S_{a,O_2}

V - the response, ie the rise in \dot{V}_E

X, Y - input and output of 1st diff. equation

x, y - input and output of 2nd diff. equation

τ_1, τ_2 - time constants of the diff. equations

G_1, G_2 - gains of equations

G, g - gains of equations when linear

I, i - measures of inhibition by output of other equation

σ_1, σ_2 - measure of saturating effect

$a \dots f$ - multipliers specifying the arrangement of inputs and outputs to the two boxes.

The parameters that define the overall relationship are a to f , τ_1 & τ_2 , G & g , I & i , and σ_1 & σ_2 . Any of these apart from the time constants may be zero.

9.3 OVERVIEW OF THE PROGRAM

9.3.1 The aim:

If the system relating \dot{V}_E to S_{a,O_2} can in fact be expressed as two differential equations, it should be possible to find an arrangement of the two equations that allows **all the data from all the subjects** to be modelled convincingly. The aim was to find the simplest such solution, ie the one that required fewest parameters to define the relationship. To achieve this it was necessary therefore:

1. To find the **arrangement** of inputs and outputs that worked best (ie consistently allowed a good fit to the data). This involved finding which out of parameters a to f could be made zero for all subjects.
2. To find which, if any, **non-linear terms** were necessary. Some terms might be needed for some subjects but not others, in which case they were kept in the final model. Other terms might not be needed for any subject, and so were excluded altogether from the model.

Once the model had been reduced to the fewest parameters possible, values for these parameters could be found for each subject. The arrangement of the equations could itself be of interest, and might suggest

possible physiological mechanisms underlying the relationship: experiments could be designed to test out such hypotheses. With some understanding of the physiological meaning of the various parameters it could be possible to classify subjects according to parameter values.

9.3.2 The Program:

The program was written in Fortran IV and assembler (Macro 11) on a PDP 11/73 mini-computer (running under TSX, a multi-terminal operating system). The program was large, using all the memory available to a terminal (and requiring the continual use of overlays), and generally the curve-fitting operation took several hours to find the best fit, an important practical consideration when designing studies. A more powerful computer system would have allowed, in the time available, the testing of more different models, a more thorough search for the best-fitting parameters for each, and a better testing of the sensitivity of fit to each parameter. The only plausible alternative to using the laboratory computer was to use a mainframe computer via a modem, but this was ruled out due to problems with data transfer and the awkwardness of using the program interactively during the model development process.

An overview of the program is as follows (summarised in Fig 9.3; full details are given in the following sections):

1. A specially prepared input file containing breath-by-breath values of S_{a,O_2} , \dot{V}_E , P_{ET,CO_2} and breath durations for a subject was read into arrays in memory. The baseline values of \dot{V}_E and S_{a,O_2} were calculated, and the S_{a,O_2} values were subtracted from the baseline (to give S in the equations listed in §9.2.2). These S values were interpolated to 1 per second to allow compensation for delays in the measurement of S_{a,O_2} , and the introduction of other delays into the system if needed.
2. A file containing values for the parameters defining the relationship was read and the parameter values displayed. These values could be amended as desired. Each parameter could be set to zero, given a fixed value or given a value that could be changed later by the program in order to find the best set of parameters.

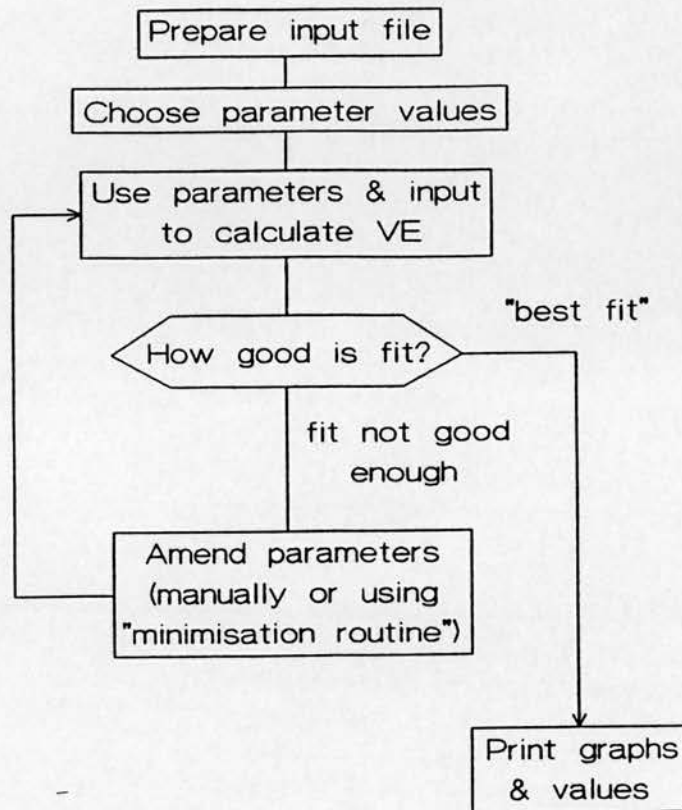


Fig 9.3 Flow diagram of the model program.

3. Using this set of parameters in the equations, the equations were solved using the values of S to obtain a set of breath-by-breath values of \dot{V}_E . These were plotted on a graphics screen alongside the values of \dot{V}_E found in the original experiment.

4. The program calculated a measure of how close the model's \dot{V}_E trace was to the observed \dot{V}_E . The parameters could then be changed manually, and steps 3 and 4 repeated in an attempt to improve the closeness of fit.

5. Alternatively, some of the parameters could be altered iteratively by the program, using a standard minimisation routine to find a set of

parameters corresponding to a minimum value of the measure of closeness of fit. This automated the process and introduced a degree of objectivity in the choice of parameters. It was possible to choose which parameters were the ones the program could manipulate, whilst keeping the others zero or at a fixed value. The choice of a different group of parameters to be varied might lead to a closer fit, and an objective method was developed to say whether one fit was statistically better than another.

6. Once the best set of parameters was found, the program allowed graphs to be plotted and the calculated points to be written into a file.

With the exception of the routines that interact with the user (which are of no scientific interest, although vitally important from the user's point of view), the various sections of the program will be described in more detail.

9.4 PREPARATION OF THE INPUT FILE

The input file of data for a subject must contain breath-by-breath values for S_{a,O_2} (the input to the equations), \dot{V}_E (to compare with the model's calculations), and the length of each breath (to allow the interpolation of the S_{a,O_2} values).

9.4.1 Reasons for preparing the file:

Frequently the step and transient data for a subject were collected separately and were therefore in separate files. This could lead to several problems which had to be overcome before the data was in a usable form for the main program:

1. The baseline S_{a,O_2} could be slightly different in the two files (usually by no more than 1%), which, if not corrected, had quite an impact on the fit. Because the stimulus was taken as the difference between S_{a,O_2} and a chosen baseline, it appeared that in one file there was always a small degree of stimulus present, even between two steps or transients.
2. The combined amount of data was large and contained a considerable amount of "uninteresting" information (during periods well away from actual stimuli where the subject was at steady state). The slowness of the

main program and the shortage of memory meant it was desirable to reduce the quantity of such data to a minimum.

3. The algorithm in the main program was designed to save computing time wherever there was a period where the S_{a,O_2} did not change. The oximeter only output integer values, and a small amount of noise on the signal could cause the S_{a,O_2} value to oscillate through one whole percent. The program interpreted each of these random oscillations away from the baseline as the imposition of a stimulus and calculated appropriate values of \dot{V}_E . If such randomness was removed during the periods where the subject was breathing air, the program worked much faster, both because it saw a longer period without change in S_{a,O_2} and because it did not have to calculate changes in \dot{V}_E .

4. Every so often many subjects took a particularly large or small breath, quite unrelated to the stimulus; and occasionally the data collection program mis-analysed where a breath should start or finish. These resulted in values for \dot{V}_E which were quite different from those of adjacent breaths, and could adversely affect the closeness of fit value.

5. Because time constants were to be calculated, it was not possible to remove data for breaths from the file completely during the periods of stimulus, as this would change the timing

9.4.2 The program to prepare input files:

A program was written in Fortran IV to carry out the processes that initially were done laboriously using the computer's screen editor. The effect of the preparation is summarised in Fig 9.4.

9.4.3 Preparation of S_{a,O_2} :

For each of 2 steps and 4 transients (or 4 steps and 6 transients) the program:

1. Extracted the 20 breaths pre-stimulus, the hypoxic period (or 5 breaths for transients), 5 breaths for the return to baseline, and 15 breaths post-stimulus.

2. Calculated the mean S_{a,O_2} from the pre-stimulus breaths, and deleted any breaths in the pre- and post-stimulus periods which lay outwith 2% S_{a,O_2} of that mean. During these periods of normoxia any large fluctuations in S_{a,O_2} were artefacts. If not deleted, these could lead to an incorrect baseline level.

3. Calculated the baseline as the mean of the remaining pre-stimulus breaths and stored this value for stage 6.

4. Zeroed all the pre-stimulus and post-stimulus values, subtracted the baseline level from the remainder (ie the hypoxic breaths and the five immediately following).

5. Concatenated all the sections prepared by stages 1 - 4.

6. Found the mean of all the calculated baselines, and added this to every value.

This procedure eliminated fluctuations in S_{a,O_2} during baseline normoxia, but did not change the data during the hypoxic section in a significant fashion (as it was the fall in S_{a,O_2} from baseline that was used in the differential equations, rather than the absolute value of S_{a,O_2} , and this was not changed).

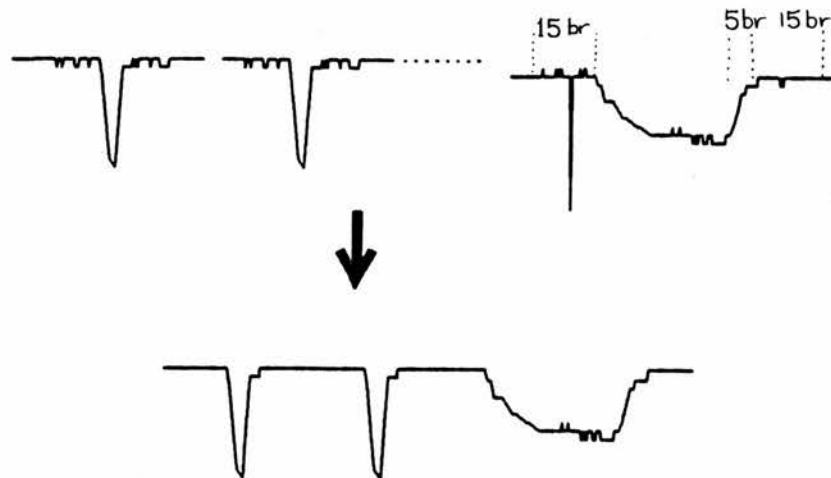


Fig 9.4 The effect of processing the S_{a,O_2} traces. Outliers, fluctuations and periods of normoxia are removed and the sections concatenated with a correction for differing baseline levels.

9.4.4 Preparation of \dot{V}_E :

\dot{V}_E was plotted for each section of the file in turn on a graphics screen, along with a list of values of \dot{V}_E , S_{a,O_2} , P_{ET,CO_2} and inspired O_2 fraction, to allow the user to mark the breaths that were clear outliers. These values were replaced with the **median** of the surrounding 5 values (including the value itself). If the value to be replaced lay within 2 breaths of either end of the section, the median of the end 5 breaths was used. The median is a more "robust" value for the replacement than the mean (Press et al, 1986 p459): the mean of 5 numbers can be heavily affected by a single outlying point, while the median will only be seriously affected if 3 breaths of the 5 are outliers, and will otherwise give a good measure of \dot{V}_E in the surrounding breaths. When the pattern of \dot{V}_E was fairly regular, the occasional very high or low value due to problems with the valve, the subject or the data collection program could easily be picked out. The problem came with the more erratic pattern of breathing found in certain subjects. Here a degree of subjectivity was inevitable, but a general principle was adopted not to change a data point if there was any possibility that it could be part of the normal breathing pattern, particularly during the hypoxic periods. Although several algorithms were tried, it appeared that this process could not be safely automated. Routines that searched for values more than a certain number of standard deviations from the mean, for example, highlighted breaths for removal at the onset of the stimulus (where in many subjects there was a sudden rise in \dot{V}_E), which was exactly the period of most interest.

9.5 THE INPUT TO THE PROGRAM

As values were read from the disk file to arrays at the start of the program, a running total of each breath duration was kept (to give the time from the start of the file as if the file were continuous), S_{a,O_2} values were subtracted from the baseline (so that all values were zero or positive), and a baseline for \dot{V}_E was calculated as the mean of all the pre-stimulus periods.

9.5.1 Interpolation of S_{a,O_2} :

The delay and time constant in the ear oximeter resulted in values for S_{a,O_2} lagging behind the corresponding values for \dot{V}_E by more than 2 seconds. To correct the S_{a,O_2} by a fixed length of time, it was necessary to interpolate the values to one per second (the shortest interval that constraints imposed by computer memory size would allow). The first three values were then omitted from the resulting array, which corresponded to a correction of three seconds: this aligned the changes in S_{a,O_2} and \dot{V}_E most closely. An error of up to half a second could be introduced here, but the time constants of interest calculated by the program were generally large in comparison, and the uncertainty in the calculations due to other causes rendered this error insignificant.

The method used to carry out the interpolation was that of Lagrange (see Chapter 5, section §5.3.1). Lagrange developed the formulae in such a way as to keep the gradient of the resulting line continuous (ie to prevent dy/dt having any sudden steps in value), so producing the smoothest curve between the given points. This resulted in a small downward curve immediately before each sudden upwards change in value (which prevented the gradient changing suddenly from zero to a large positive value). Other smoothing occurred where the S_{a,O_2} would otherwise have had sharp changes in gradient. Given that the S_{a,O_2} data was only specified originally as an average value per breath (which itself implies a considerable degree of smoothing), the interpolation smoothing was insignificant.

9.6 SOLVING THE DIFFERENTIAL EQUATIONS

The complexity of the equations meant that, except for a few mathematically trivial cases, they could not be solved algebraically (ie it was not possible to find an equation in the form $y = f(x)$ which did not include any differential terms). The output curve therefore had to be found numerically. There are several standard methods for this, one of the most popular and useful being the Runge-Kutta method, of which the fourth-order form is generally used as being the most efficient (in terms of precision per number of calculations).

9.6.1 The Fourth Order Runge-Kutta Method for Integration:

(Press et al, 1986, pp547-577). This is a way of integrating equations of the form

$$dy/dx = f(x,y) .$$

In our case there were two equations to be solved of the form

$$dy/dt = f(t,y) .$$

The function f depended on time through a dependence on the system's input, which was changing with time.

The Runge-Kutta method works in a step-by-step way, estimating the next value of y from the current values of y and t , using only gradients as calculated from the original equation (rather than using higher order differentials as some other methods do - this is one of the advantages of this method). If the current value of y is y_i and of t is t_i , then the next value of y when t has increased by a short period h to t_i+h is given by:

$$y_{i+1} = y_i + (C_1 + 2C_2 + 2C_3 + C_4)/6$$

where

$$C_1 = h.f(t_i, y_i)$$

$$C_2 = h.f(t_i+h/2, y_i + C_1/2)$$

$$C_3 = h.f(t_i+h/2, y_i + C_2/2)$$

$$C_4 = h.f(t_i+h, y_i+C_3) .$$

The accuracy of this new value is to the order of h^4 .

9.6.2 Adaptive step size:

Because the accuracy of each estimate depends so heavily on the size of step, h , this must be chosen carefully. Clearly greater precision is gained when h is small, but this must be offset against the greater number of calculations required to cover a given range of t . When y is changing rapidly, a slightly inaccurate estimate will lead to a wrong starting point

for the next estimate, leading to an even more inaccurate next estimate, and so on, so the inaccuracy will tend to become magnified rapidly. In such cases h must be very small. When y is changing slowly the gradients calculated will be small, and the difference made to the estimate of the next y by a small amount of inaccuracy will barely be noticeable. There is no reason for h to be small, and so for efficiency h should be increased so that fewer calculations are needed.

A standard routine (Press et al, 1986, p558) for choosing the most appropriate step size for any particular part of the curve uses the following algorithm.

1. The user specifies a degree of accuracy that each estimate must be made to.
2. The first step size is a guess.
3. This step size is used to calculate the next value of y .
4. This value of y is re-calculated using two steps, each of half the step size, which will result in a more accurate estimate.
5. The two values of y are compared. If the difference is greater than the accuracy specified in step 1, the step size is again halved and the middle point of the two-step estimate in step 4 is re-calculated. If the difference between these two values again falls outside the degree of accuracy specified, the step size is again halved, and this process is repeated until the accuracy reaches the desired limit (or until h becomes so small that it is not changing significantly between calculations, in which case the routine produces an error message).
6. When values of y calculated in one step and two half steps agree within the required accuracy, this value of y is taken as the new starting point, and h is increased according to a formula which takes into account the closeness of agreement. When y is changing slowly, so that the accuracy required is usually achieved on the first trial, h rapidly increases in length.

9.6.3 Changes required to this routine:

The adaptive step-size algorithm results in calculated values of y which are grouped tightly together where y is changing rapidly and spaced out where y is more stationary. The program, however, required a value for y at pre-determined times, ie the end of each breath. Values for times other than these were not stored once they had been used to calculate the following values. The step-size routine had to be altered so that if the next step was longer than the time remaining before the next breath, the step was set to be the remaining time (so that a point was calculated exactly at the time of the next breath), and the step size that should have been used, according to the algorithm, was used for the following step.

This of course reduced the efficiency of the step-size controller by introducing shorter steps than would otherwise have been used just before the end of each breath period, and by restricting the maximum step-size to being less than the length of a breath. Time was saved, however, by performing a check on the values of the input of each differential equation for the current and next breaths, and on the outputs for the last and current breaths: if these did not change by more than a very small amount then the next value for the output had to be effectively the same as the current one, and was therefore taken as such without any calculation. Significant amounts of computing time were saved by this method during periods of steady state (such as the normoxic periods separating the stimuli).

9.6.4 Using the Runge-Kutta routines in the program:

Fig 9.5 summarises how the routines fitted together in the program. At the lowest level was the routine that calculated the gradients dy/dt and dY/dt using the differential equations with a given set of parameter values, and the current value of S , the fall in S_{a,O_2} . This routine was called repeatedly by the Runge-Kutta routine in calculating the values of Y and y , and the Runge-Kutta routine itself was called by the adaptive step-size routine which presented it with the current time and the step size to use. The routine at the highest level was one that requested values for Y and y for each breath and used them to calculate the response V . This routine also kept a running total of the difference between the observed and calculated

responses to use in the closeness of fit measurement, and plotted graphs on the graphics screen if desired.

Each time this highest-level routine was called to obtain a response curve of values (for perhaps 700 breaths), the Runge-Kutta routines calculated values for Y and y for many thousands of intermediate values. Each time the Runge-Kutta method was invoked, 4 gradients needed to be calculated, and this was by far the most time-consuming part of the method (the Runge-Kutta calculations themselves being very simple). During curve-fitting many hundreds of response curves were calculated, so the gradient-calculating routine were called millions of times. Great pains were taken to ensure that this routine was written in the most efficient way possible, and once the routine had been shown to work without error it was converted to assembler, and the outputs of the assembler and Fortran routines were compared exhaustively over the range of inputs to ensure that no errors entered during conversion.

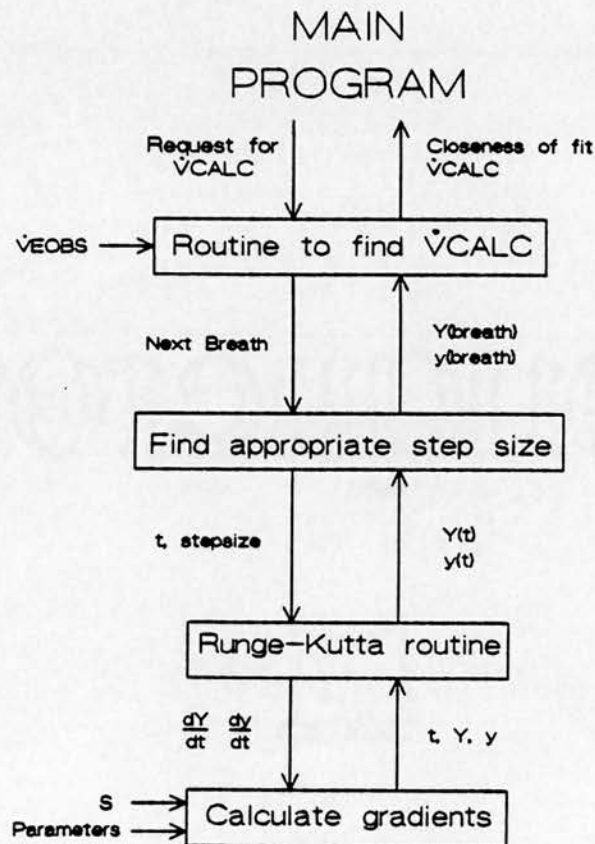


Fig 9.5 Summary of the Runge-Kutta routines showing parameters passed between the routines. Routines towards the bottom of the diagram are called more frequently than those higher up.

9.7 OBTAINING A VALUE FOR CLOSENESS OF FIT

[The phrase "closeness of fit" is used in preference to the more usual "goodness of fit" because the latter has specific statistical meanings, not all of which are applicable here.]

A standard measure used in curve-fitting for how closely a calculated curve fits an observed one is given by summing the squared difference between points on the two curves, ie finding

$$\sum (\dot{V}_{E\text{obs}} - \dot{V}_{E\text{calc}})^2.$$

The smaller this value, the better the fit between calculated and observed curves, **assuming that the observed values are distributed approximately normally about their running mean.** In fact, this value is often of use even when the observed values are not normally distributed, though the strict mathematical theory supporting its use cannot be used. It is a convenient and easily-calculated value which generally results in fits which seem adequate by most standards (though when the assumption of normality is false it is quite difficult to establish objective standards to decide exactly what is meant by the best fit). It is least useful when the normal distribution is quite clearly distorted, the most common example being when there are outliers (occasional observed values that lie well away from the rest, such as those in our experiments due to swallowing or sighs or the valve sticking momentarily). The sum of squared differences value is usually calculated repeatedly, changing parameters of the equations giving the calculated values until the sum of squared differences reaches a minimum (though perhaps its most common occurrence is in the "Least Squares" method for linear regression where the simplicity of the situation allows the minimum to be reached via an algebraic rather than iterative route).

Unfortunately the sum of squared differences could not be used exactly as above. There were two problems:

1. Very many more values of response (and therefore of squared difference) were calculated for steps than for transients. This was simply due to the greater number of breaths involved (for example, a study

involving 4 steps and 6 transients may have 240 breaths during the steps and only 30 during the transients). A bad fit to the transients therefore made only a relatively small increase in the total sum of squared differences, whereas a slightly imperfect fit to the steps had much more impact. The steps would always be fitted well, to the detriment of the transients, after minimising the standard sum of squared differences. The transients, however, were of equal importance in the current investigation.

2. Values of the response between the stimuli (once the response had returned to baseline) were set equal to the baseline value. This value had been calculated as the mean of the normoxic sections of the observed response, so the curve automatically lay along the best fit to these sections. Using these values as part of the sum of squared differences calculation would add nothing that would affect the fit, but rather reduced the sensitivity of the total sum to imperfections in the fit during the interesting sections of curve. This would not, of course, ultimately affect the fit once the minimum value was found, but it would affect the statistical analysis that was done later using that value.

The solutions to these problems were simple. The sum of squared differences was weighted so that transients had the same impact on the overall sum as the steps. The weighting factor, by which all the differences during the steps were divided, was given by

$$W = \frac{\text{total number of breaths in file during steps}}{\text{total number of breaths in file during transients}}.$$

In the example given above where a file contains 240 breaths of steps and 30 of transients (counted as the number of breaths where the S_{a,O_2} is not at its baseline), the weighting factor would be 240/30.

The second problem was solved by using only points where the calculated response was above the baseline. A potential problem with this, though, was that if the set of parameters caused the response to remain at baseline throughout, then the closeness of fit value was zero simply because no values were used in the sum. In such circumstances the program set the fit value to a very large number, so that such fits would be rejected. It was possible for the calculated response to be above zero for a few breaths and

then return immediately to baseline for the rest of the fit, resulting in a fit value which was the sum of a very few points: this could occur with inhibition, for example, and would lead to a lower fit value than a closer fit involving more points. To avoid this, the sum of weighted squared differences was divided by the weighted number of points used, N_W , where

$$N_W = N_T + (N_S / W)$$

N_T = No. of non-baseline values of \dot{V}_{Ecalc} for Transients

N_S = No. of non-baseline values of \dot{V}_{Ecalc} for Steps

W = weighting factor as before.

Omitting to weight N_S would have led to a bias towards a fit where N_T was small and steps were fitted well, because of the use of weighting in the fit calculation.

An important advantage of this division by the weighted number of points was that the closeness of fit value produced was directly comparable between fits to different data sets for one subject even when the numbers of steps and transients used were different. Unfortunately, a large and variable part of the closeness of fit value depends upon the breath-by-breath variability in the subject's observed \dot{V}_E , and this could vary markedly between individuals: the closeness of fit value could not therefore be compared directly between subjects.

There are still potential pitfalls in this choice of measure of fit, but it is highly unlikely that any would arise with the sort of equations and inputs used. A short period of near perfect fit (eg for the first three breaths out of a step), with the remainder of the curve at basewline level, would result in a very low fit value, even though most of the curve was not being fitted at all. This is only barely possible: while the input to the equations is non-zero it is very unlikely that the baseline value would be output unless the gains were small, the time constants large and probably some form of inhibition was occurring, all of which would lead to at least a few badly fitting points before baseline was reached - which should increase the fit value sufficiently to indicate a bad fit. More likely, but less of a problem, is that a slow tail-off after the end of the stimulus (where a slow time

constant caused the response to fall gradually towards the baseline) could increase the number of points to be divided by without adding very much to the sum of squared differences, and so might improve the fit value for an otherwise poor fit. This would lead to a potential favouring of long time constants. In fact, as will be described in the next chapter, the response following the return of S_{a,O_2} to its baseline value was artificially truncated rapidly to baseline, for other reasons, thereby also eliminating this problem.

9.7.1 Summary of measure chosen for closeness of fit:

The value calculated for each response curve was given by

$$\text{Fit} = \frac{\sum (\dot{V}_{E\text{obs}(T)} - \dot{V}_{E\text{calc}(T)})^2 + \sum (\dot{V}_{E\text{obs}(S)} - \dot{V}_{E\text{calc}(S)})^2}{N_W} / W$$

where

$$N_W = N_T + N_S / W$$

$$\text{and } W = \frac{\text{total number of breaths in file during steps}}{\text{total number of breaths in file during transients}}.$$

$\dot{V}_{E\text{obs}}$ is the observed \dot{V}_E and $\dot{V}_{E\text{calc}}$ the calculated \dot{V}_E (during steps (S) and transients (T)) where the calculated response was above the baseline value. This fit value may be compared directly between fits for a given subject, but not between fits for different subjects.

9.7.2 A further use for weighting:

Although not used for most of the fitting in this thesis, an option was built into the program to force closer fitting to the beginning of steps that started with one or more breaths of a very hypoxic gas mixture (as in the experiments of Chapter 6). Studies of this sort lay emphasis on what happens initially in the response, but the fitting process fits all points without favour, leading to the same problem as with fitting transients and steps. If requested, this option doubled the weighting for the first 5 breaths of every step, making it less advantageous for the fitting process to fit the end of the step well at the expense of a poor fit to the beginning.

9.8 THE MINIMISATION ROUTINE

Curve-fitting is an iterative process. A set of parameters is used, the response curve calculated for the input, and then the parameters are changed in an attempt to improve the response curve. Changing the parameters can be done manually each time, but it is simpler and more objective to use one of the standard methods for changing the parameter values within a program. Such "minimisation routines" use well-tested algorithms to seek a set of parameters which result in a minimum value for a given function (the function here being the closeness of fit value described above). There are many standard algorithms, each with advantages and disadvantages. The method chosen for this work was the **Simplex** method. The algorithm will be described briefly, the modifications that had to be made to adapt it to the given problem will be detailed, and then potential problems associated with minimisation will be mentioned.

9.8.1 The Simplex Method:

(Press et al, 1986, pp289-293). This method, developed by Nelder and Mead (1965), was chosen for two main reasons. Firstly, it required no differentiation of the function to be minimised (which would have been awkward and time-consuming, as analytical differentiation with respect to each parameter was not feasible given the flexible arrangement of the differential equations, and numerical differentiation would therefore have had to be used). Secondly, the algorithm was easy to understand, to visualise (to a degree) and to program. This last factor is probably the most important. The routine had to be heavily adapted to allow flexibility in the differential equations, and having a clear and simple routine to work from reduced the likelihood of errors being made, and therefore reduced programming time. The method is not, however, as efficient as some others (ie on average it makes more calls to the time-consuming closeness-of-fit calculation routine on the way to the minimum than those others do). The final program was therefore slower than it might have been, but as it could be left running at night when time was not of the essence, the trade-in for faster programming was worthwhile.

The method works by considering a multi-dimensional shape (a "simplex") with one more vertex (and dimension) than the number of parameters being manipulated. Each vertex corresponds to a set of parameter values, and so a closeness-of-fit value can be calculated for each (which, for simplicity, will be referred to as "the function value" for the rest of this section). At each iteration, the vertex associated with the highest function value is moved to a new position chosen according to simple rules. If a new position has a lower function value then it is used for the new vertex, and the routine goes on to replace the next highest vertex. The rules for moving the vertices cause the simplex to move and stretch out towards regions of low function value, contracting in to a point when it reaches a minimum.

The process is easily visualised when there are only two parameters to be manipulated. If the two parameters are each given a cartesian axis, any combination of the two values will have a unique position on a plane, and for each position there will be an associated function value. The function value can therefore be plotted as a three dimensional surface above the plane. The simplex for the two parameters will be a triangle with each of its points on the function surface. The method will cause this triangle to change shape repeatedly so that it moves down into a valley and finally contracts round the lowest point in that valley. This three-dimensional example will be returned to when discussing problems associated with minimisation routines.

The simplex routine is initialised by providing it with a value and a step length for each parameter. The step lengths are used to specify the positions of the vertices of the initial simplex, and the first few moves of vertices will be of the order of their associated step lengths (therefore small initial step lengths will result in a small simplex moving slowly at first, while large step lengths will result in a simplex initially of a very large size making large movements).

9.8.2 Adapting the routine for the program:

Several adaptations to the standard routine were required:

1. The main model equations were specified by a large number of parameters, and one of the aims of the investigation was to decide which parameters were important, which were unchanging between subjects and which could always be left at zero. There had therefore to be a way for the user to communicate to the program which of the parameters were to be manipulated, so that the simplex routine could create and use a simplex based only on those parameters. The step length was used to flag which parameters were to be used in a given run: the user set the step length associated with any parameter that was to be kept constant to zero before starting the fitting process, and the simplex routine used only parameters which had associated non-zero step lengths, ignoring the others.

2. Computers can only deal with numbers up to a fixed maximum value, and any higher number causes an error to be registered. The highest number allowed on the PDP/11 was of the order of 10^{34} . Certain combinations of parameters led to a fit so bad that the closeness of fit value exceeded that number. An error trapping procedure had therefore to be built into the minimisation and gradient-calculating routines, with an overflow check after every calculation which could conceivably result in an overflow. As soon as an overflow was detected, the closeness of fit function was set to the value 9.9×10^{33} .

3. Problems can occur if two or more vertices have the same function value, and this was quite likely to occur at the start while the simplex was large, as several of its vertices might have the overflow function value associated with them. If all the vertices have this same overflow value, the method would generally fail completely, so it was essential that minimisation was always started with a set of parameters associated with a non-overflowing function (so that at least one vertex was smaller than the others). The simplex method always works on the vertex with highest function value, and the standard routine, when there is more than one vertex with the same high value, will always start with the first on the list, and will not proceed to the next until a lower value has been found for the current one. This was undesirable, as the frequency of overflows meant that parameters higher up on the list would generally be favoured in the fitting process (leaving those lower down until the fine-tuning stage). Early on there is nothing to suggest that one parameter is intrinsically

more important than another, so the routine was modified to choose randomly any of the overflow vertices to work on, rather than simply the first.

4. The simplex has a tendency to stop when it finds a "local minimum" (this is discussed in the next section). To make sure that the minimum found was an important one, the simplex was always re-formed with this minimum used as one vertex, and the minimisation was set going again: if the minimum was important, the simplex rapidly returned to this same point; otherwise it went on to find a new, better, minimum (which itself had to be checked by re-starting the simplex). When two consecutive minimums were identical, the program was stopped, and the associated parameters were taken as giving the "best fit".

9.8.3 Problems with minimisation routines:

Minimisation routines are designed to find a minimum, ie a set of parameters where making a small change to any parameter results in an increase in the value of the function. It is quite possible that a function has more than one minimum, but the minimisation routine will only find one of them, and that will not necessarily be the one with absolutely the smallest function value. Returning to the example of a two parameter function where the function values can be plotted as a surface, the minimisation routines will hunt for the floor of a valley, but could quite easily find a small depression (a "local minimum") where the function increases in all directions (for a short distance at least), and would stop as this would fulfil the criteria they are looking for. If the surface is shaped like an eggbox, there may be many valleys of similar importance: without trying all of them, the routine would not be able to say whether it had found the deepest one.

That is the problem: the routines only try a limited number of points on the surface, and if none of these are on the slopes of the deepest valley then that valley will not be noticed. The points chosen depend upon the starting place, and starting from a different position can result in a better (or worse) set of parameters chosen as a "best fit". There is therefore an element of chance in the proceedings, as the starting position is generally chosen arbitrarily. Mapping the whole surface by finding function values at regular points throughout the range of all the parameters is not

practicable when there are many parameters, due to the immense calculation time needed. It is necessary just to accept that the minimum found is simply a very good fit, and is the "best fit" only in the sense that it is the best fit in the region of space tried.

The problem of local minima is less acute when few parameters and relatively simple equations are used, as the hypersurface (the many-dimensional analogue to the three-dimensional surface of the example) is likely to curve quite smoothly. As the number of parameters and the complexity increases, so does the potential for the hypersurface to acquire folds and troughs where chance interactions between parameter values cause a local decrease in the function value. If this were not the case, the ideal approach to the problem under investigation (seeing which of the total set of parameters were important) would be to allow all the parameters to be manipulated in finding a best fit, and then to see with which parameters the function value changed most rapidly, and which could be made zero without significantly affecting the fit. This was the approach planned before the minimisation routine was written; but it soon became clear that manipulating such a large number of parameters was intolerably slow, and the routine kept stopping at local minima with function values well above those found by manipulating a small subset of the parameters. The approach was therefore abandoned.

9.8.4 Problems specific to the simplex method:

Problems occurred with the simplex routine that might not have occurred with other minimisation algorithms. These were not of great consequence. Occasionally, by chance, a vertex of the simplex would be moved so that it finished exactly on a line joining two other vertices. The simplex thus effectively lost a dimension and became unable to find a minimum as the hypersurface now had more dimensions than it had. The symptoms of this were usually an endless repetition of a series of manoeuvres without any improvement in function value. This problem was dealt with by re-starting the simplex at a different point.

Certain (uncommon) shapes of surfaces could cause the simplex to oscillate about a point, again repeating endlessly a series of manoeuvres, the function value at one vertex being low, and at the others increasing and decreasing during the cycle. One such shape appeared to be a

minimum with a lip all around it. One vertex lay within the dip, but as the simplex contracted around this lowest point, the other vertices increased their function values, and each in turn was reduced by moving it **away** from the minimum, increasing the size of the simplex which then had to contract again, and so on. This problem could usually be dealt with by stopping the routine, taking the best vertex as the point of best fit and re-starting the simplex from there as with other minima: this is very unlikely to produce the same points that gave rise to the oscillation.

Perhaps the most important criticism of the simplex method is that of speed. It worked well and efficiently when the number of parameters being manipulated was less than four. The time taken to find a minimum increased rapidly with the number of parameters thereafter (though this was in part due to the greater complexity of the surface being searched). A three parameter fit typically took 20 minutes to find a minimum. A five parameter fit could take several hours.

9.9 A METHOD FOR SHOWING THAT ONE FIT IS STATISTICALLY BETTER THAN ANOTHER

Using five parameters might result in a lower closeness of fit value than using four parameters, but it is important to be able to say whether this decrease is significant. The data that was being fitted, the subject's \dot{V}_E , had a degree of randomness in its pattern, and a slightly better fit value might be obtained by having, say, a slight rise in the calculated response where the observed response rose, simply due to the randomness. The underlying pattern would not have been fitted better, only the random pattern. The problem was to distinguish between fitting the randomness and fitting the true underlying response, which is a problem analogous to that solved by Analysis of Variance (which checks whether differences between two samples are genuine or due only to the randomness in the sampling).

The best fit to a set of data cannot fit exactly due to the randomness of the observed \dot{V}_E . The difference between the best-fitting curve and the observed \dot{V}_E can be used to assess the size of the random element. The sum of squared differences between observed and best fit \dot{V}_E is referred to as the **Residual Sum of Squares (RSS)**. The points where the calculated \dot{V}_E

is at the baseline value must be excluded from the sum: these points are constrained to a value that has been calculated in a different way from the fitting process. Because weighting was used in the fitting process, the RSS must also be calculated with this weighting. It can be seen that the RSS is very closely related to the closeness of fit value (which was one of the reasons why this measure of fit was chosen), and indeed it is simply the value associated with the best fit before the division by the number of points.

The RSS is calculated in a similar way to a variance (which is the sum of squared differences from a mean value). This means that the F-test can be used to compare one RSS with another. If the RSS gained using q parameters is RSS_q (using n points in the calculation), and the RSS gained by adding another parameter is RSS_{q+1} , then Snedecor's F is given by

$$F = \frac{(RSS_q - RSS_{q+1}) / 1}{n - (q+1)}$$

with 1 and $n - (q+1)$ degrees of freedom. The numerator is the improvement in fit, and this is compared against the randomness in \hat{V}_E which is measured by the RSS of the better fit, RSS_{q+1} , in the denominator. The number of degrees of freedom associated with each RSS is given by the number of values in the sum, n (which assumes that all these differences are independent of each other), minus the number of parameters (q or $q+1$). The number of degrees of freedom associated with the numerator is therefore 1, as

$$(n - q) - (n - (q+1)) = 1.$$

If the value of F exceeds the 5% value for those degrees of freedom, then the improvement in fit is considered significant. If it does not, then there is no justification for using the extra parameter in the fitting process.

9.10 THE GENERAL APPROACH TO USING THE PROGRAM

There were many possible models contained within the differential equations of the program. The problem was to find a model structure applicable to all the data - that is, to work out which parameters could always be kept zero or at a fixed value independent of which subject's data was being fitted. It was not possible to try every possible combination of

parameters on all the data for every subject: this would take a prohibitively long time. Before starting to use the program, therefore, the following strategy was designed to get round this problem:

1. Work initially on the data from one subject. The final model must work for **all** subjects, so any models that do not work for **this** particular subject have logically failed the test. It is sensible to choose a subject whose response is not at the extreme end of the range.

2. Prepare a representative sample of data for this subject, ie choose a few of the transients and steps. A similar argument to that in 1. applies: the model should fit all the transients and steps and therefore can be rejected if it does not fit a subset of them.

3. Try to find the simplest structure that fits this data. This stage will be largely a matter of trial and error, guided by a knowledge of what each element in the equations can do. As the aim is the simplest model possible, one way of proceeding is to make as many parameters as possible zero and see whether an approximate fit can be obtained with the remainder, and then to add in more and more parameters to see which improve the fit.

4. Once what appears to be a good fit is obtained, the same set of parameters should be applied, unchanged, to more data from the same subject. The fit should still be reasonable. If not, changes should be made as appropriate, and stage 3. repeated.

5. Once the fit is good for all that subject's data, each of the parameters used should be changed slightly both upwards and downwards to see how sensitive the fit value is to these changes. If the fit value is not at all sensitive to changes in a parameter, then it may be possible to remove that parameter from the structure (ie to set it to a constant value) as it probably is not contributing to the fitting process, and so is not a value that can be defined precisely by a subject's data. If, on the other hand, the fit value is highly sensitive to a parameter (changing wildly when the parameter is only changed by a small amount), then consideration should be given to rejecting use of that parameter, as anything that affects the response so dramatically is unlikely to reflect a physiological control mechanism.

6. The structure obtained should next be applied to other individuals. The non-zero parameters will, of course, vary from subject to subject. It is likely that the structure will need to be adapted to cope with the new subjects, but care must be taken to leave it capable of fitting the original subject. The changes to the structure will hopefully be slight: if not, and especially if the changes degrade the fit to the first subject, step 3. should be returned to.

7. Once one successful model is found, others should be sought. It is possible that more than one model structure could fit the data equally well.

8. Each model should be validated (Murray-Smith & Carson, 1988), a process that to some extent occurs concurrently with the earlier stages. The model structure should be checked to see that it is internally consistent, plausible and useful.

This strategy was followed, leading to the development of a model fulfilling the criteria.

CHAPTER 10

THE DEVELOPMENT OF A SUCCESSFUL MATHEMATICAL MODEL

10.1 INTRODUCTION

This chapter details how, using the program described in the last chapter, a model was developed capable of fitting almost all the data from the original study.

Following the strategy detailed at the end of the last chapter, data from a single subject was used in the early development of the model, and several examples of best fits from models tried en route to the final one will be given for this subject. The model was expanded with reference to the other subjects, and plots of best fits for the data of every subject using the final model are presented. Although this chapter might give the impression of a straightforward logical path to the final model, it will be appreciated that the actual route was much more complicated, with time-consuming detours up blind alleys (of which only a few examples are mentioned here).

10.1.1 Adequacy of fit:

Frequently in this chapter fits to the data are described as adequate or inadequate. The last chapter explained why standard statistical methods could not be used to give an **absolute** measure of how good a fit was (§9.7), so descriptions of the adequacy of a fit depend heavily on visual inspection and will therefore contain an element of subjectivity. An inadequate fit failed on one or more of the following criteria:

1. The fitted curve should follow the general shape and magnitude of the observed response during the steps, and should therefore always be approximately the mean of the surrounding observed values. If the fitted curve lay outside the random noise on the observed response for most of a step the fit was almost certainly inadequate.

2. The observed transient response should not be systematically over- or under-estimated.
3. The rate of increase of response at the start of hypoxia should be similar to that observed.

Despite the unavoidable subjectivity in this, it was found that different people would generally arrive independently at the same conclusion over adequacy of fit. As described in §9.7, the assessment of **improvement** in fit was wholly objective.

Occasionally the relationship between \dot{V}_E and S_{a,O_2} changed noticeably part way through a hypoxic step, particularly towards the end. This might take the form of a reduction in \dot{V}_E once both \dot{V}_E and S_{a,O_2} had reached temporarily a plateau; the S_{a,O_2} subsequently would fall due to the decreased oxygen intake (see, for example, the data for Subject 1 (2nd and 3rd steps, Fig 4.3). Clearly another mechanism to the simple stimulus-response relationship built into the model was occurring, perhaps the first signs of central hypoxic depression (see §1.2.5), and so the model could not be expected to fit these regions accurately. The calculated response increased due to the greater fall in S_{a,O_2} , and therefore diverged from the decreasing observed response, failing the criteria for adequacy of fit set out above. Such periods were therefore ignored when assessing adequacy.

10.2 THE DEVELOPMENT OF A MODEL USING DATA FROM ONE SUBJECT

The subject chosen for the development of the model was Subject 10. The magnitude of her response was in the middle of the range of the ten subjects, with neither the exaggerated response of some nor the seemingly reversed response of others (whose response to transients was greater than that to steps). Her data offered the additional advantage of being very well defined, the breath-by-breath randomness in \dot{V}_E being small in magnitude compared with the underlying response (ie the signal-to-noise ratio was high). The advantages of this were that a bad fit would not be masked by the noise, and that any improvements in fit would be clearly seen (and any that were seen were likely also to be statistically significant according to the test described in the previous chapter, as the denominator in the

calculation of F would be small). A model developed using data as good as this might not be useful for data where the parameters were less well defined; but that was something that could only be established once a model had been decided upon.

A file containing two transients and two steps (one at 15% and one at 12% O_2) was prepared as described in the last chapter. Initially fitting was carried out using only one of the differential equations, to check both the program and the work done in Chapter 8. As expected, no adequate fit could be found. The other equation was then brought into play, and numerous arrangements of two linear differential equations were tried. Once it was seen which arrangements were working well, the other available parameters (describing the non-linear effects) were used to see which, if any, improved the fit.

10.2.1 One differential equation:

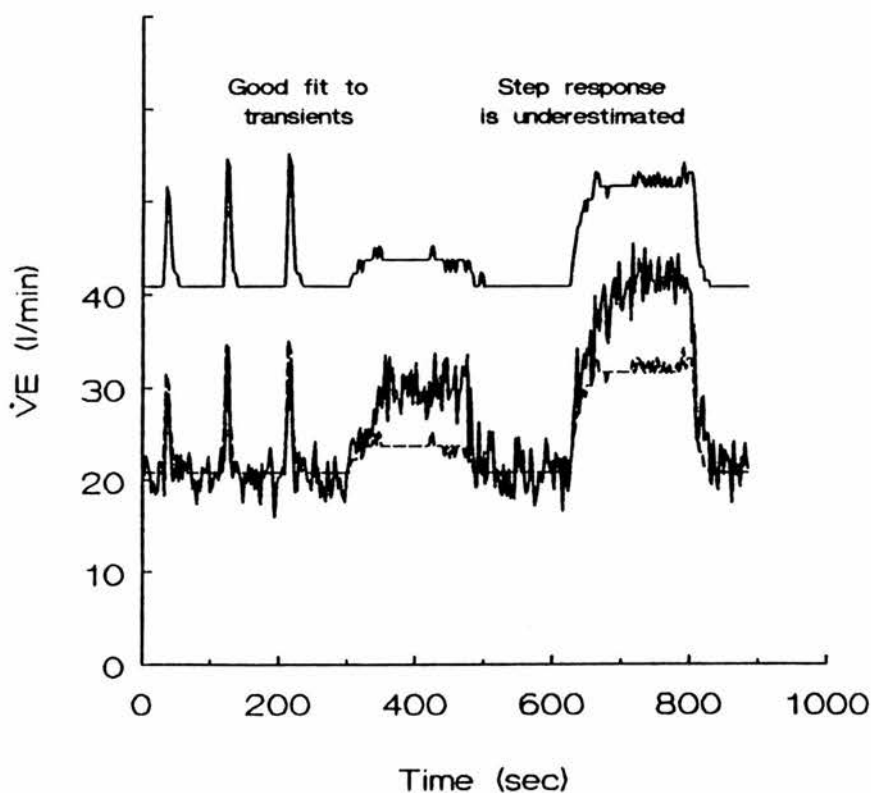


Fig 10.1 Fit using a single linear differential equation. The calculated \dot{V}_E is shown as a broken line and is also repeated at an offset of 20 l/min for clarity.

In Chapter 8 a theoretical argument was used to show that one differential equation with reasonable non-linear terms was inadequate for fitting the sort of response found. This was checked, by setting the gain of the second equation to zero and making the overall response equal to the output of the first equation.

The best fit obtained with effectively a single linear differential equation is shown in Fig 10.1. In this and subsequent figures, the calculated response is shown as a broken line superimposed on the observed response, and is also repeated for clarity offset by 20 l/min. It can be seen from this figure that the steps were underestimated by the model while the transients were overestimated, and the tail-off following the end of each stimulus while the calculated \dot{V}_E returned to baseline lasted longer than with the observed response. It was clear that one differential equation was, as was concluded in earlier chapters, quite inadequate.

10.2.2 Two linear differential equations:

The program was written to allow the input to each differential equation to be any combination of S (the input to the whole system) and the output of the other equation (§9.2.2). To arrive at a simple model structure, the simplest arrangements were tried first. These were the arrangements where the input to an equation was **either** S **or** the output of the other equation (but not a combination of the two). The set of possibilities was further reduced as at least one of the equations must have S as its input (or the system would not be reacting to the stimulus at all), and the output of each equation must be used in some way, either as an input to the other equation or as part of the final output (otherwise the equation serves no purpose). With these restrictions, nothing was gained by having values other than 1 and 0 for each of the parameters a-f. Only five arrangements fulfilled these criteria, with parameter values as follows:

<u>a</u>	<u>b</u>	<u>c</u>	<u>d</u>	<u>e</u>	<u>f</u>	
1	0	0	1	0	1	serial
0	1	1	0	1	0	serial
1	0	1	0	1	1	parallel
1	0	0	1	1	1	hybrid
0	1	1	0	1	1	hybrid

The first two are equivalent - the order of the equations being swapped round - and the same applies to the last two. These arrangements are shown in Fig 10.2.

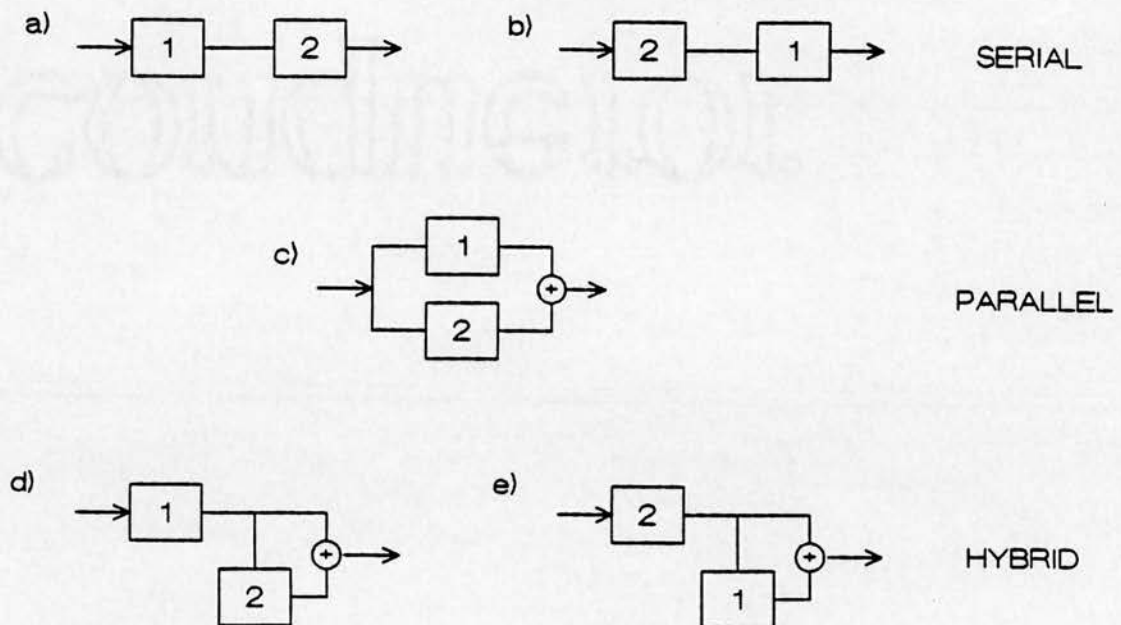


Fig 10.2 The five different arrangements of two differential equations.

The serial arrangement failed to improve the fit obtained from one differential equation. This is easy to explain. If the first equation has a fast time constant, then its output will be very similar to S , but magnified by the gain. The second equation will thus have as its input something almost identical in shape to S , and so its output will be very similar in shape to the output had it used S directly. The first equation is not making

much difference to the final shape of response. A similar argument applies when the second equation has a fast time constant: the overall shape of the response will be left much the same as the output from the first equation. The only possibility left is when both equations have slow time constants, and then the fast initial response during transients cannot be modelled. At least one of the equations must have a fast time constant for this.

The parallel arrangement produced a very much improved fit (both visually and statistically) when one equation had a fast time constant, the other a slower one (Fig 10.3). Because the output from both equations was used in the final response, the "fast" equation allowed the initial rapid rise in \dot{V}_E of the transients to be fitted, whilst the "slow" equation increased the eventual maximum response during steps.

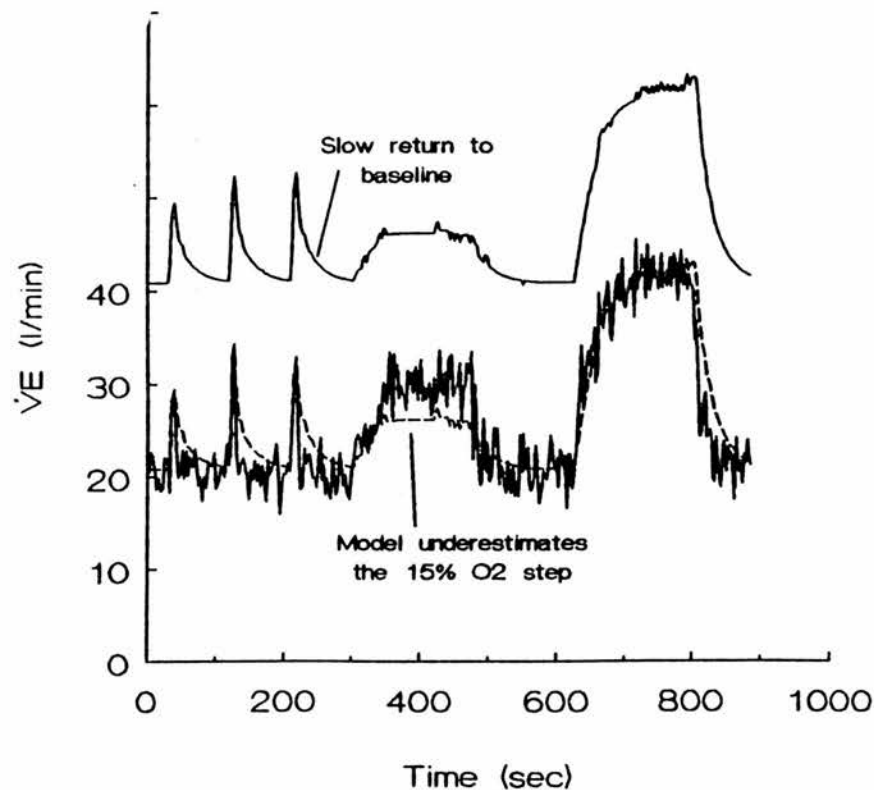


Fig 10.3 Fit using two differential equations in parallel, one with a slow time constant, the other fast.

The **hybrid arrangement**, where the output from one equation acted as input to the other as well as itself forming part of the output, allowed a fit almost identical to the parallel arrangement when the first equation had a fast time constant. This was to be expected: the fast time constant leaves the shape of the output of the equation very similar to the shape of input S (though magnified by the gain), so the second (slow time constant) equation is seeing effectively the same shape of input as the slow equation in the parallel arrangement. If the product of the gains of the two equations in the hybrid arrangement is equal to the gain of the slow equation in the parallel arrangement, and the gain of the fast equation is the same, then the overall outputs of the two arrangements will be nearly identical. The fit when the slow time constant equation acted first in this arrangement was not good.

There seemed, therefore, to be two potentially useful structures around which to base a model, both involving two equations, one with a fast time constant, the other with a slower one, both equations contributing to the overall response (Fig 10.4). Because the two structures are so closely related (as the value of the faster time constant is only a few seconds) the following sections refer only to the parallel arrangement, though everything applies equally well to the hybrid arrangement.

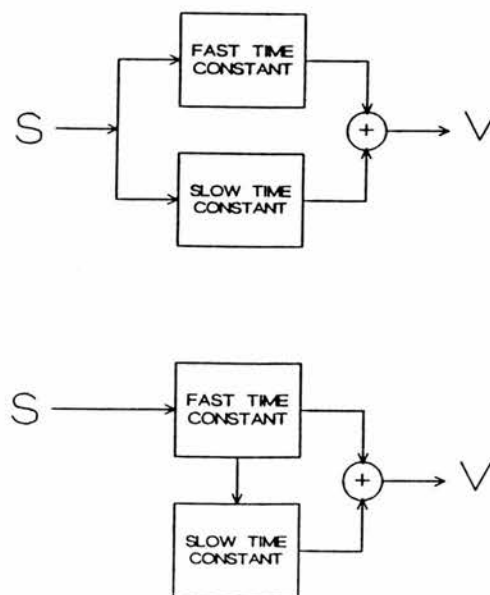


Fig 10.4 The "Parallel" (upper) and "Hybrid" (lower) arrangements.

10.2.3 Shortcomings of the parallel arrangement:

The response calculated with the parallel arrangement differed from the observed response in two main areas. Firstly, the return to baseline ventilation following the end of the stimulus was more gradual in the calculated response than in the observed response, both for steps and transients. This was due to the slow time constant equation. Secondly, although it was possible to obtain a reasonable fit to the transients and the 12% O₂ steps, with the same parameter values the calculated response to the 15% step was noticeably underestimated (Fig 10.3). Similarly, a fit to transients and the 15% step overestimated the response to 12% O₂.

Small additions to the parallel arrangement model allowed these differences to be overcome convincingly:

10.2.4 Cut-off of the output of the slow differential equation:

The slow time constant needed to give the difference in size of response to transients and steps meant that the output of the slow differential equation took several seconds to return approximately to zero after the end of the stimulus. This period, however, corresponded to a period of hypocapnia in the subject due in the transients to the increased ventilation of the preceding breaths, and in the steps to the termination of the compensatory flow of CO₂ (no attempt to maintain isocapnicity was made at the end of steps in these experiments). The subject's \dot{V}_E therefore fell sharply to baseline (or below) at the end of the stimulus due to hypocapnic depression of ventilation. S_{a,O_2} , however, did not return immediately to its baseline, but remained 1 or 2 % below for a few breaths, a consequence of the reduced ventilation. The model sees this as a continuing small stimulus, and this compounds the problem of the slow time constant tail-off.

The discrepancy between \dot{V}_{Ecalc} and \dot{V}_{Eobs} during those few breaths adversely affected the "best fit" found by the model. The program minimised the measure of closeness of fit by underestimating the main part of the transient or step in order to reduce the difference between observed and calculated ventilation during the tail-off period. This

compromise by the fitting process resulted in gains which were artificially small in the region of most interest.

Several methods for removing the tail-off were tried. The most obvious, limiting the size of the slow time constant, prevented a good fit to both steps and transients and so was not the solution. If the output of the two differential equations was multiplied rather than added, then at the end of the period of hypoxia as the fast time constant equation's output fell rapidly towards zero, it forced the product of the two outputs towards zero despite the tail-off of the slow equation. By this means the tail end of steps and transients could be fitted well (Fig 10.5), but the method was abandoned as the first few breaths of steps were fitted very poorly, as here the slow equation's output kept the overall output low, preventing the sharp rise in response seen in the observed \dot{V}_E . The only method tried that proved successful was a simple one: the output of the slow differential equation was set to zero when the S_{a,O_2} increased to within 2% of the baseline value (i.e. when S fell below 2).

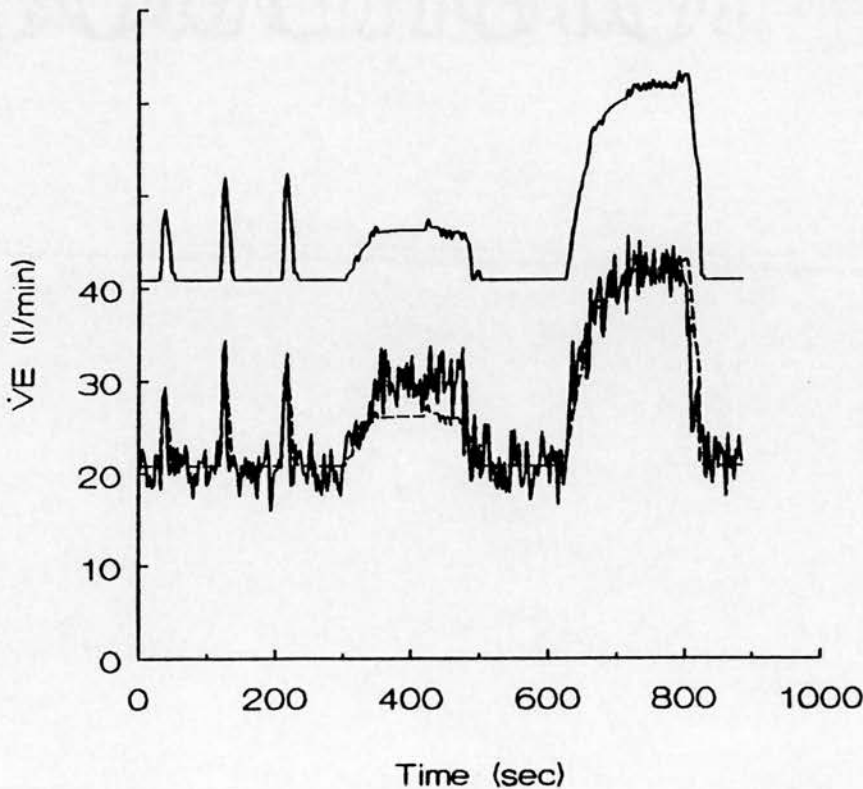


Fig 10.5 Fit using two linear differential equations with cut-off. The slow time constant equation's output is truncated following hypoxia when $S < 2$.

This approach might seem artificial and without physiological basis, but a case can be made that it was neither of these. The cause of the sudden fall in ventilation observed following the end of a step or transient was almost certainly due to hypocapnia at the carotid body reducing the respiratory drive. This was due to a change external to the simple \dot{V}_E - S_{a,O_2} system on which the model was derived. The model as it stood assumed that S_{a,O_2} was the only stimulus affecting \dot{V}_E and so could not be expected to model periods where a second stimulus had effect. The truncation of the output to zero as the input approached zero was effectively a crude model of the hypocapnic effects. Rather than spend time developing a more sophisticated model for this period (which would be of limited use due to the complexity of a two-stimulus system represented by very few data points), the simple approach was adopted and this complex area was effectively removed from the analysis.

10.2.5 Addition of the "Saturating Effect":

The second problem with the simple parallel model was that the response to the 15% O_2 step was underestimated. This could be improved by use of the "saturating effect". As described in section §8.3.1 the "saturating effect" reduces the gain of the differential equation so that as the input increases, the output is affected less and less until a maximum possible output level is reached which cannot be exceeded, no matter how great the input. When this term was used in the **slow** differential equation, a significantly better fit was obtained, using the test of significance described in Chapter 9, §9.7 (Fig 10.6). With this addition, the parallel arrangement model can be represented mathematically as:

$$\tau_1 \cdot \frac{dY_1}{dt} + Y_1 = G_1 \cdot S$$

$$\tau_2 \cdot \frac{dY_2}{dt} + Y_2 = \frac{G_2}{(1+\sigma \cdot S)} S \quad \text{if } S > 2$$

$$\underline{\text{or}} \quad = \quad 0 \quad \text{if } S < 2$$

$$V = Y_1 + Y_2$$

$$(\dot{V}_{E\text{calc}} = V + \dot{V}_E(\text{normoxia}) ; \quad S = S_{a,O_2}(\text{normoxia}) - S_{a,O_2})$$

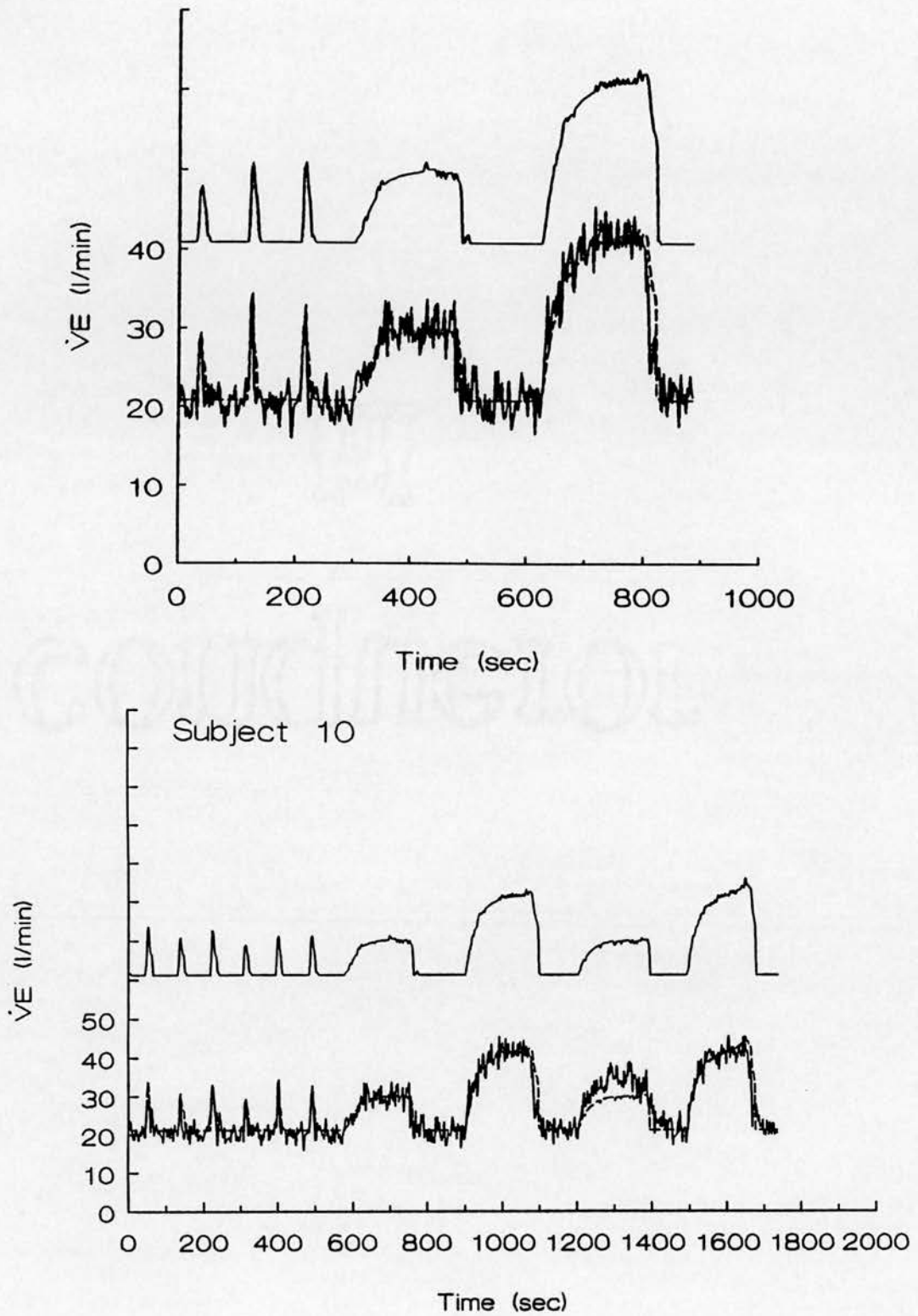


Fig 10.6 Fit using two differential equations in parallel with saturating effect. The lower graph shows the fit to all the transients and steps for this subject. The model output is repeated at an offset of 40 l/min.

This model fitted the data for the transients and three of the four steps convincingly (Fig 10.6). Possible reasons for the discrepancy between calculated and observed values of the third step will be discussed in the next chapter.

10.3 APPLICATION OF THIS MODEL TO OTHER SUBJECTS

The aim was to find a model that fitted all the data of all the subjects. In the last sections many possible models were excluded by showing that they did not fit one of the subjects (and therefore logically could not fit **all** of the subjects), and a simple model was found that could not be excluded on the basis of the data of this one subject. The next step was to test the model on data from other subjects, and then to accept, reject or amend it as appropriate. If amendments were necessary, the model must, of course, be re-applied to the data of the first subject to make sure that it remained compatible.

Several of the subjects (1,2,4,5,8,10) were fitted reasonably well by this model. It became apparent that the **time constant of the faster equation** was always very much less than one second, and little improvement in fit could be gained by trying to find a precise value for it, while the attempt took much computational time. It was therefore set arbitrarily to **0.1 seconds**. This may seem unrealistic physiologically, but it must be remembered that the stimulus being reacted to by the model (S_{a,O_2} as measured by the oximeter) already contained an instrument time constant of about 2 seconds. The model's short time constant merely meant that the equivalent time constant in the subject was of a similar size to that of the oximeter and could not therefore be measured with accuracy. All that could be said with certainty was that the subject's time constant was less than 3 seconds, and probably less than 2. The value 0.1 seconds was chosen for convenience: anything shorter added considerably to computational time (as the steeper slopes required many more steps in the Runge-Kutta algorithm), while above 0.2 seconds the model could not mirror some of the rapid rises in observed \dot{V}_E in some subjects, especially during transients.

Some subjects could not be fitted well by this model. In Subjects 7 and 9 a fit to the steps considerably underestimated the response to the transients,

while in Subjects 3 and 6 the opposite was the case. Rather than reject the model outright (and lose the good fit to the other subjects) an attempt was made to add to the model further terms that would improve the fit to these subjects. With data for so few subjects available it was important to avoid the situation where the model developed was so complicated that in effect each subject was fitted by a separate model. Adding a new term to the equation for each subject would make the model worthless as the whole point was to find one structure that was universally applicable. However, the fact that more than half of the subjects could be fitted reasonably by the model already described seemed to testify to its value as a basis for further adaptation, and it was clear from even a cursory glance at the data of subjects 7 and 9 that their responses were very different from those of the other subjects (having a low response to step stimuli but high to transients (see Fig 4.2g and 4.2i in Chapter 4)). It therefore seemed reasonable to extend the model rather than reject it, so long as the changes did not add too much to the overall complexity. The need for checking the resulting model on new subjects of course increases with these extensions.

10.3.1 "Inhibition":

Subjects 7 and 9, unlike all the others, showed a reduced response to step hypoxia compared with their response to transients. If, as before, it was assumed that the rise in \dot{V}_E was a response to the fall in S_{a,O_2} , then clearly either the response to the transients was being exaggerated by some mechanism, or the response to steps was being repressed. The latter seems more likely: if the response depended only on the level to which the S_{a,O_2} had fallen then it would seem unlikely that any mechanism that acted over a short period to increase response to transients was not also noticeable over a longer period, or at least over the first few breaths. How was the control mechanism to know that the hypoxia was to be short-lived in one instance and not in another?

One potential mechanism for increasing the response to transients was that if the rate of fall of S_{a,O_2} was being detected and was itself acting as part of the stimulus. This idea was considered and rejected in Chapter 6 for subjects with a greater response to steps than transients, but might be applicable here. The rate of fall of S_{a,O_2} during the plateau section of a step, however, was practically zero. The rate of change part of the stimulus

would no longer be acting and so the response to a step should be highest during the first few breaths and should then diminish. This was not borne out by the data. The failure to see this cannot be blamed on a slow time constant smoothing out the effect, as the rapid rise of \dot{V}_E during transients implies a fast time constant. Rate of fall would therefore have to be built into a more complicated hypothesis: perhaps it only acted as a stimulus above a certain threshold; or perhaps its effect died away only after S_{a,O_2} returned to baseline. These two hypotheses might be verified or refuted by the experiment, described in Chapter 6, involving different rates of fall of S_{a,O_2} leading to a step (the fastest rate of fall should produce the highest \dot{V}_E plateau), but unfortunately these two subjects were not available to take part in this study. These hypotheses, however, seem rather complicated, and we were seeking a simple solution.

It seemed much more likely that it was the response to steps that was being inhibited in these subjects. This was backed up by their generally low responses to steps (compared with most other subjects). A plausible hypothesis involves an inhibitory mechanism which comes into play after hypoxia has been detected: so long as the initial rate of fall of S_{a,O_2} is slow enough for the inhibition to take effect before the \dot{V}_E has had a chance to increase to its (uninhibited) maximum, the response to step hypoxia will be low throughout, while the sudden fall of S_{a,O_2} during a transient causes a sudden rise in \dot{V}_E before the inhibition is sufficiently advanced to prevent it. This hypothesis implies two things about the inhibitory mechanism: that it is activated in some way by the hypoxic stimulus (otherwise it would be present already before at least some of the transients, leading to a reduced response); and that it does not achieve full strength immediately hypoxia is detected (or again the transient response would be inhibited immediately). Again, it is unfortunate that the two subjects did not take part in the rates of change of S_{a,O_2} study (Chapter 6) as this would have provided a good check of this hypothesis. If true, the response to a step starting with a rapid fall in S_{a,O_2} should have a sharp initial peak before inhibition set in.

The hypothesis was examined by building it into the model. In the model inhibition effectively means the reduction of one or both of the gains. The level of inhibition must depend on a process that has the fall in S_{a,O_2} as an input but, as was shown above, does not achieve full strength

immediately that S_{a,O_2} begins to decrease. This could be built into the model as a third differential equation, but a simpler solution was to use the slow time-constant differential equation. Indeed, the possibility of inhibition or potentiation between the two differential equations was one of the attractions of moving from one to two equations.

The simplest arrangement of this, and therefore the one that was tested first, was for the output of the slow equation, Y_2 , to reduce the gain of the faster equation, G_1 , in a linear fashion. This was achieved by multiplying G_1 by the factor $(1-(I.Y_2))$ where I is a parameter representing the degree of inhibition. I , in this arrangement, will have a value much less than 1, and will be greater where inhibition is more severe, and zero when there is no inhibition.

One consequence of having one equation inhibit the other was that the "hybrid" arrangement of differential equations, which up to this point had been as effective as the "parallel" arrangement, became unworkable. The "hybrid" arrangement had the output of equation 1 as the input to equation 2; equation 2 now affected the output of equation 1 and a feedback loop was formed. The overall output was often oscillatory, and sometimes became unstable. If this form of inhibition was to be included in the model, the "hybrid" arrangement had to be rejected.

Fig 10.8g (p174) shows the output of the model including inhibition for the data of Subject 7. The inclusion of the inhibition term made a statistically significant improvement to the fit, and so the model could not be rejected outright, but the fit was clearly not perfect. The transients with the greatest observed response (eg the 5th and 7th) were still underestimated by the model, though most of the rest were fitted well. The calculated response rose more rapidly at the start of the third step than the observed response. Unfortunately the large size of the random element in this subject's \dot{V}_E during the steps prevented a detailed examination of more complex inhibitory effects as the signal-to-noise ratio was too low to differentiate significantly between alternative models.

The fit to the data of Subject 9 using this model was considerably worse (Fig 10.8i), especially for the transients. The model was able to fit reasonably well a subset of the data (3 of the transients and 2 of the steps), the fit being comparable (visually) to that for Subject 7, but the parameter

values changed considerably with the fit to the entire data set, suggesting that the parameters found initially were not very robust. Examination of the data (Fig 4.2i) shows that the model was faced with an impossible task. The first 3 transients all had nearly identical falls in S_{a,O_2} , but the associated changes in \dot{V}_E were, respectively, 16, 22 and 4 l/min. There was also a marked variation in the responses to steps, with the 3rd step clearly showing a greater response than the 1st (both were 15% O_2 steps), and the 4th showing a considerably reduced response compared with the 2nd (the 12% O_2 steps). The signal-to-noise ratio was low throughout. The "best" fit to the data was therefore a compromise, and so appears to have failed throughout.

The only tentative conclusion to be gained from this would be that inhibition in some form was **possibly** the cause of the unusual response in these subjects, and that this **may** be because the process represented by the second differential equation was inhibiting the process represented by the first. A more complex inhibitory arrangement cannot, however, be ruled out without obtaining more data from these subjects, in particular data involving different rates of fall in S_{a,O_2} and different degrees of hypoxia.

10.3.2 "Potentiation":

The fit in Subjects 3 and 6 suffered from the opposite problem to the one described in the last section. Here the response to steps was very much greater than that to transients (by a factor of 6 in Subject 3). The subjects' maximum response to step hypoxia was higher than the model's best fits. A good fit to the steps could only be achieved by letting the model grossly overestimate the response to transients. It was clear that something in addition to processes already in the model was acting to increase \dot{V}_E during the steps for these individuals. Rather than admit another independent stimulus, it was realised that "potentiation" (the opposite of inhibition) could occur between one differential equation and the other. For instance the output of equation 1 could increase the gain of equation 2, or equation 2 that of equation 1. These two possibilities were tested in the model using a method similar to that for inhibition. The gain of an equation was multiplied by $(1+(P.Y))$ where P is a measure of the degree of potentiation and Y the output of the other equation. The fit was significantly improved, and was better when the output of equation 2

potentiated the gain of equation 1 than when the potentiation was in the other direction. This meant that potentiation and inhibition could be combined to use a single parameter, I , where positive values caused inhibition and negative values potentiation. It cannot be inferred from this that physiological mechanisms behind the potentiation and inhibition are necessarily different aspects of the same thing (though this is an interesting possibility), but combining them into one parameter simplifies the calculation of parameter values and makes the model more convenient to use.

10.4 THE FINAL MODEL

The simplest model that allowed a reasonable fit to be made to all ten subjects was as follows:

$$\begin{aligned}\tau_1 \cdot \frac{dY_1}{dt} + Y_1 &= (1 - I \cdot Y_2) \cdot G_1 \cdot S \\ \tau_2 \cdot \frac{dY_2}{dt} + Y_2 &= \frac{G_2}{(1 + \sigma \cdot S)} S & \text{if } S > 2 \\ &= 0 & \text{if } S < 2 \\ V &= Y_1 + Y_2\end{aligned}$$

$$(\dot{V}_{E\text{calc}} = V + \dot{V}_E(\text{normoxia}) ; S = S_{a,O_2}(\text{normoxia}) - S_{a,O_2})$$

This is shown schematically in Fig 10.7, and plots of the best fits for all the subjects (as described in the next section) are given in Fig 10.8a-j.

Up to five parameters (G_1, G_2, τ_2, σ and I) summarise the response of a subject to both step and transient hypoxia. In some subjects the response could be fitted with $\sigma=0$ or $I=0$. The same results as this model might be achieved with a model of greater complexity and more parameters, and with more data the form of the "inhibition" might be refined, but this model was built up carefully to fulfil the initial criteria of simplicity and universality.

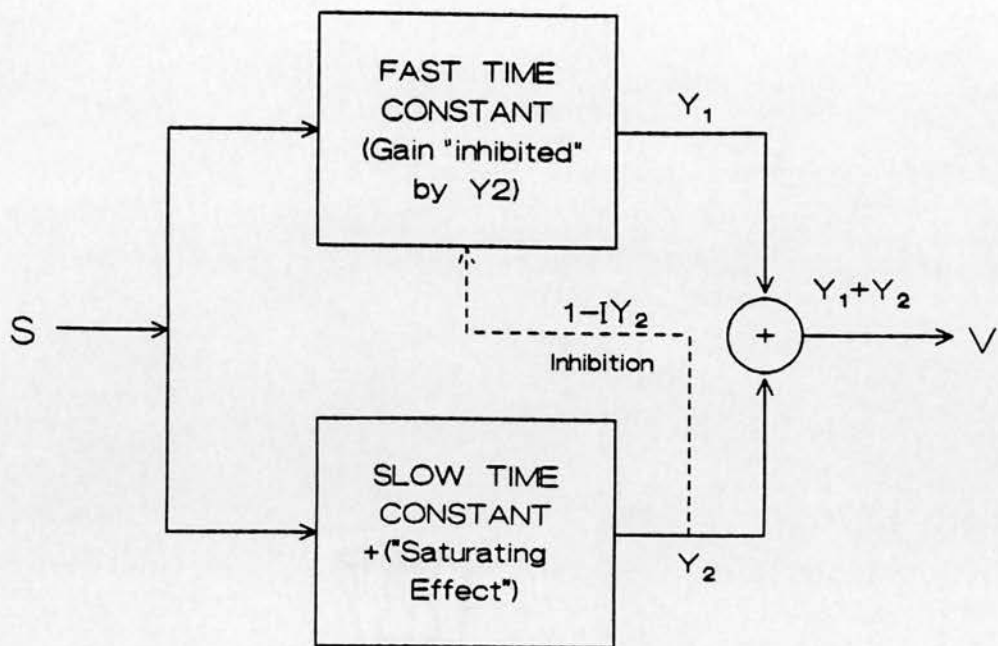


Fig 10.7 *The final model. A schematic representation of the final model equations. Each box represents a differential equation, each taking S as input. The non-linear effects are marked, but were unnecessary in some subjects.*

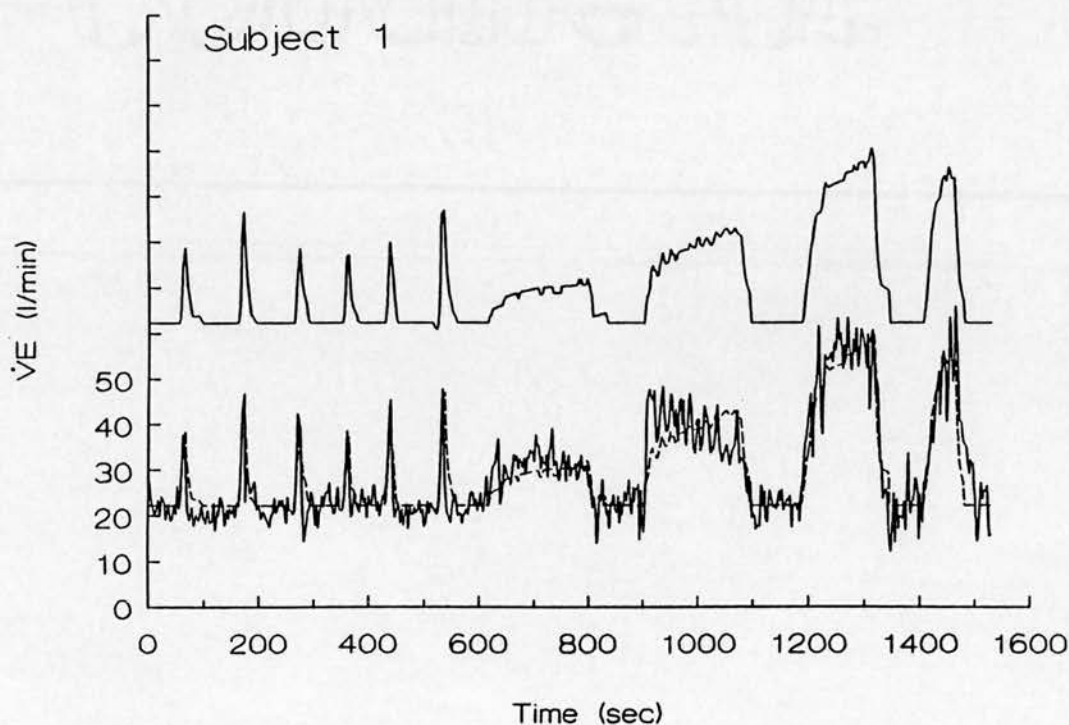
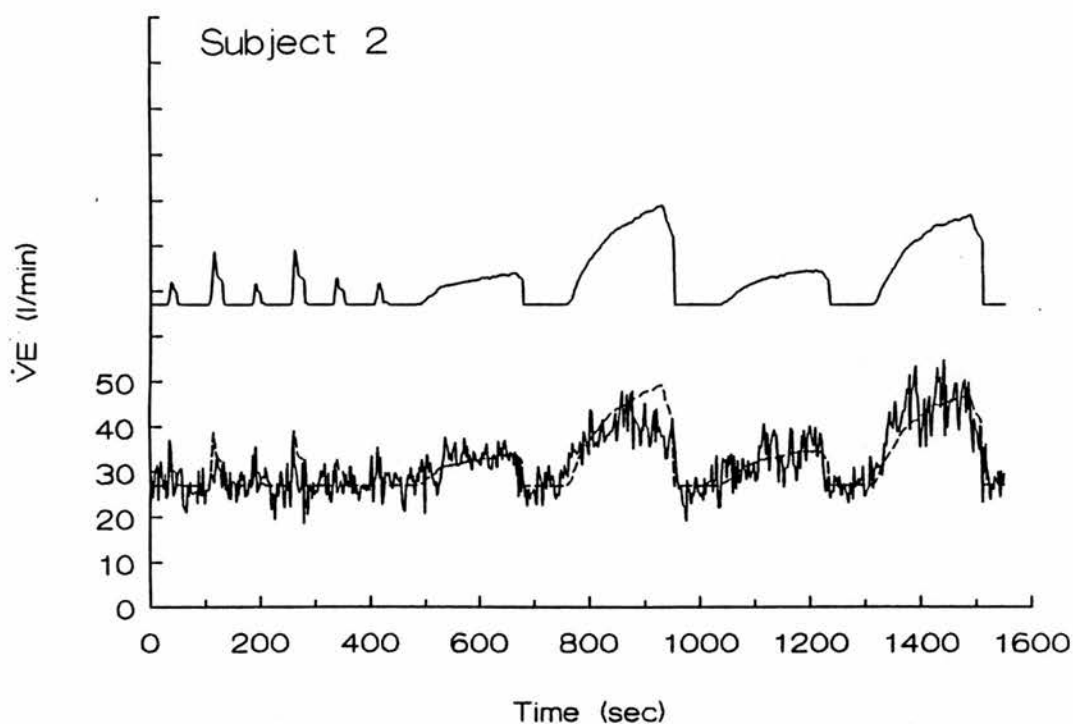
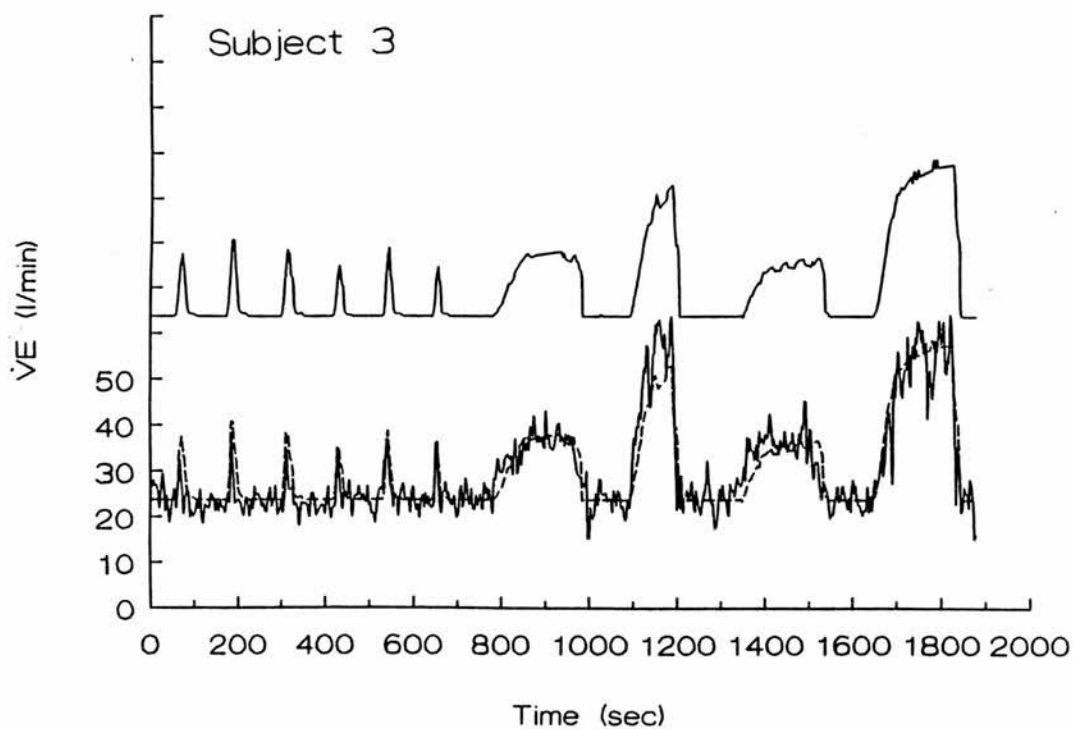


Fig 10.8 *Plots of "best fits" for all ten subjects (continued over the next few pages). The non-linear effects required for each fit are indicated. Parameter values are given later in Table 10.1.*

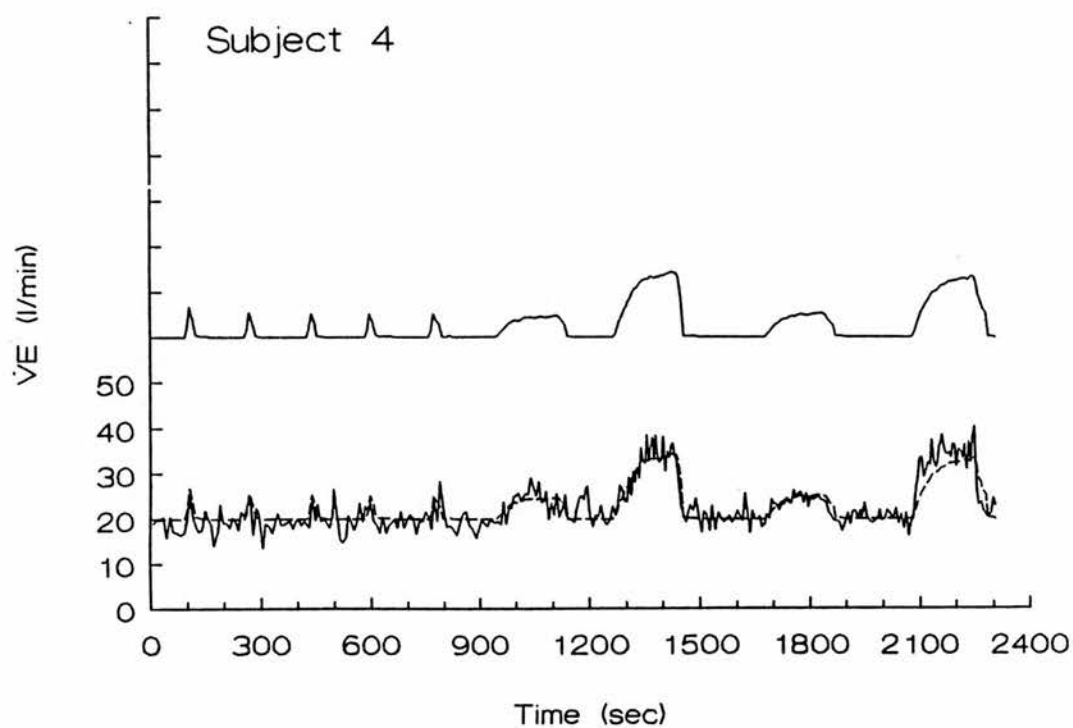
10.8 a) *Subject 1 - saturating effect and potentiation.*



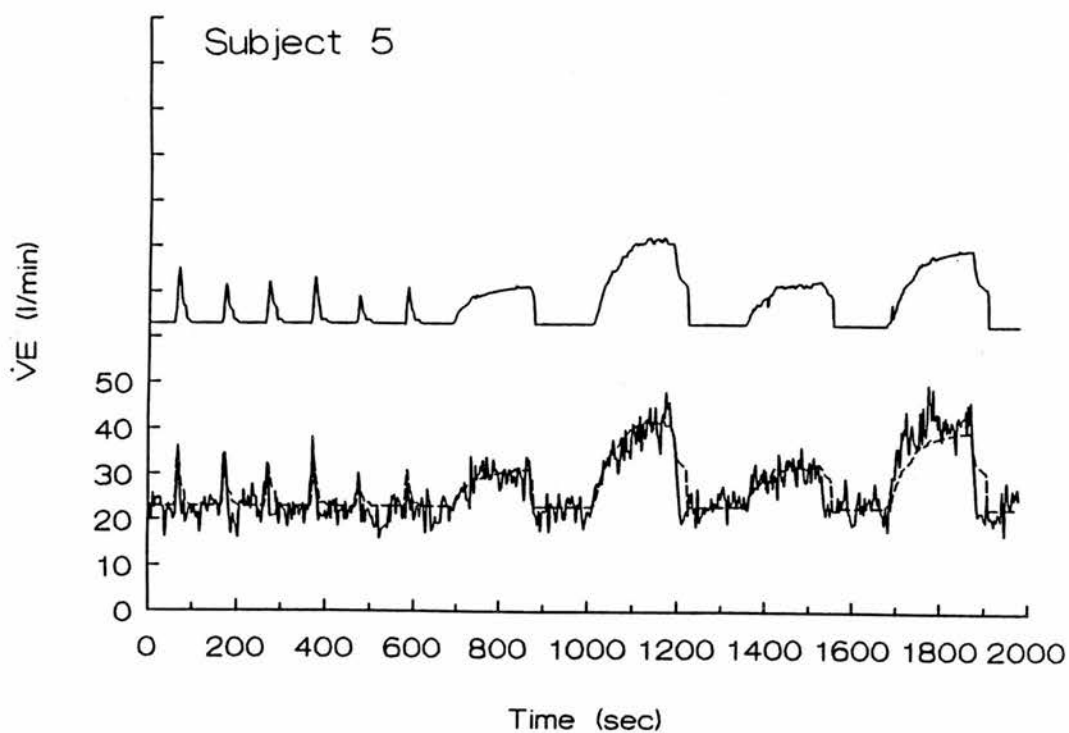
10.8 b) Subject 2 - linear



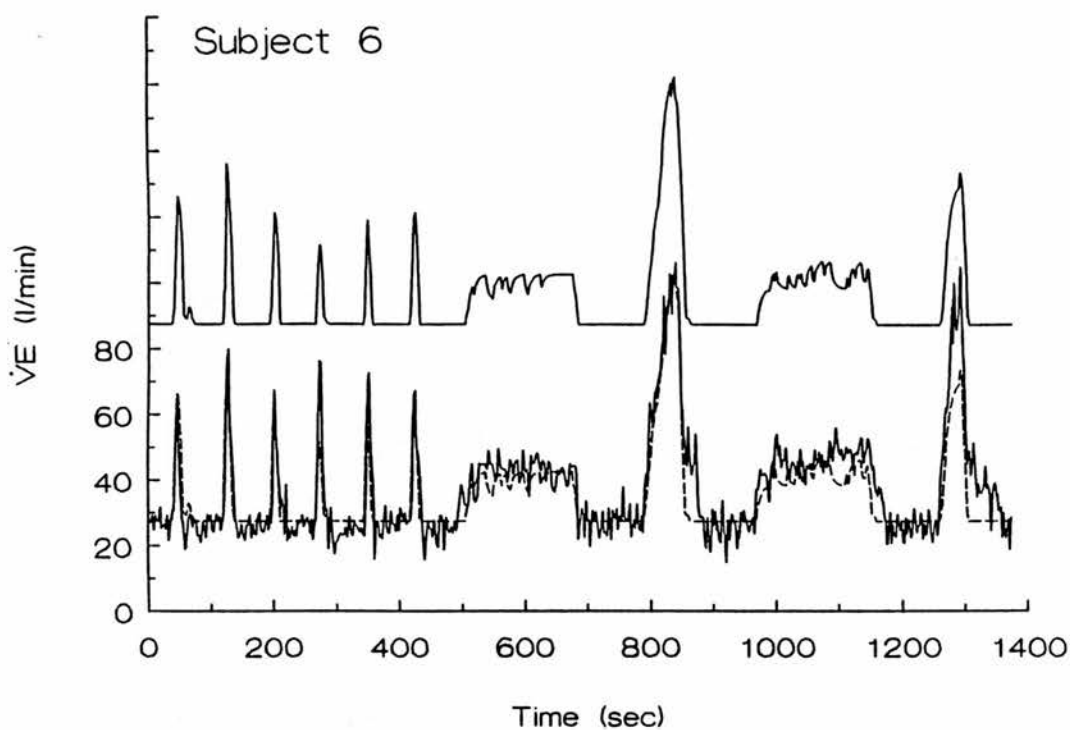
10.8 c) Subject 3 - potentiation



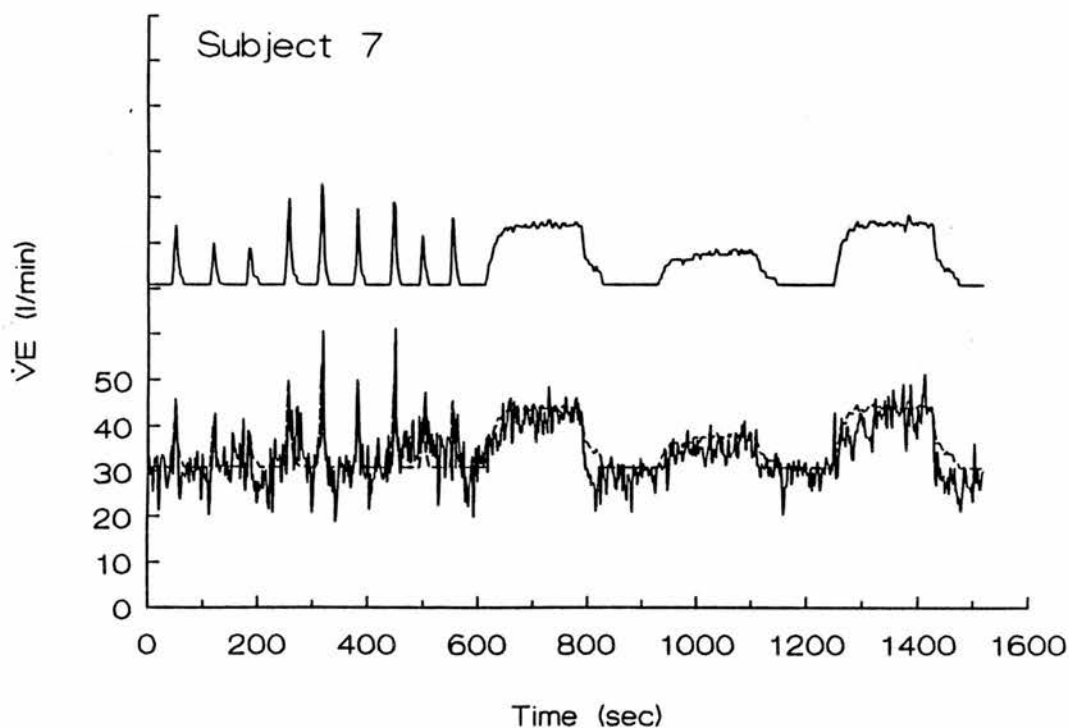
10.8 d) Subject 4 - linear



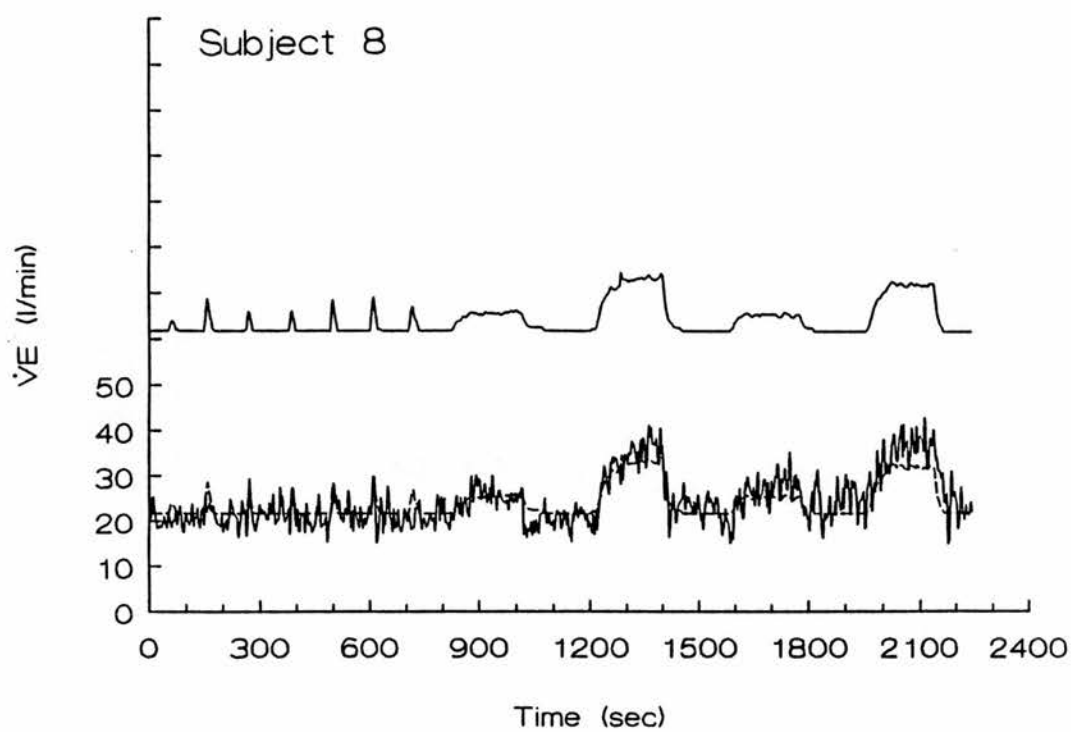
10.8 e) Subject 5 - saturating effect



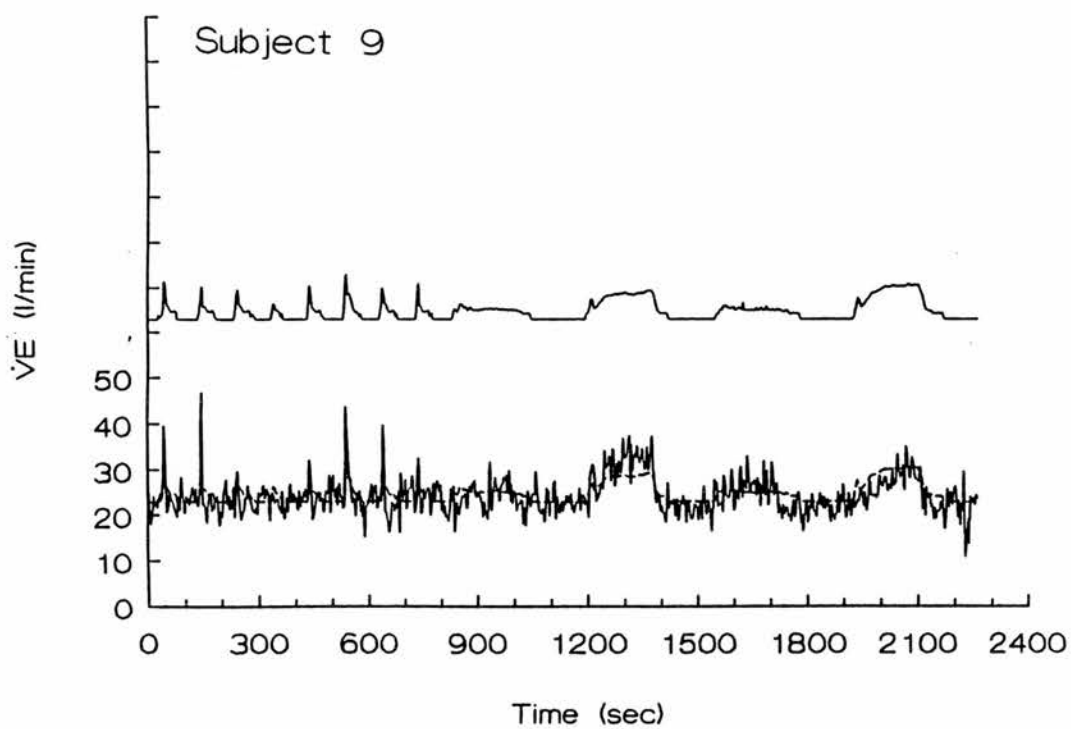
10.8 f) Subject 6 - potentiation



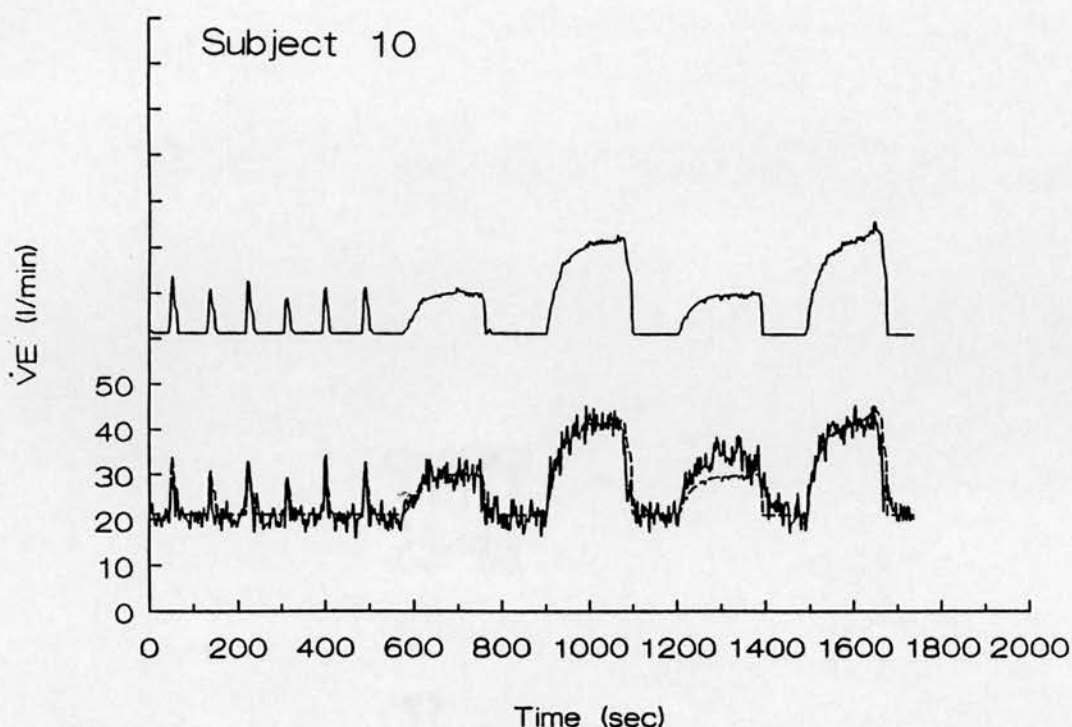
10.8 g) Subject 7 - inhibition



10.8 h) Subject 8 - linear



10.8 i) Subject 9 - inhibition



10.8 j) Subject 10 - saturating effect

10.5 PARAMETER ESTIMATION FOR EACH SUBJECT

10.5.1 The program:

Having settled on a final model, a program was prepared to find values for the five parameters as efficiently as possible. The program worked in the same way as that used for developing the model, except that much of the flexibility of the earlier program was now redundant and was removed to improve performance, and the parameters other than I were prevented from taking negative values. To allow the program to run unsupervised, all the input needed from the operator was obtained before starting the fit, and the process of repeatedly re-starting the minimisation until consecutive best fits did not show a significant improvement was automated. A hard copy of the best fit curve was printed automatically once the program finished.

10.5.2 The method: The data of each of the ten subjects was re-fitted using the new program. First the fit was made using just parameters G_1 , G_2 and

τ_2 . The fit was then repeated using σ as well (ie including the "saturating" effect) to obtain a new closeness of fit, and again with I (allowing "inhibition" or "potentiation"), and with both σ and I. If a fit was significantly better than the previous one then the new set of parameters was adopted, otherwise the simpler set was chosen.

10.5.3 Results:

Values for the parameters giving the best fit for the 10 subjects are shown in Table 10.1 (below) and are plotted in Fig 10.8 (above).

Fits to the data of Subjects 2, 4 and 8 were not improved by the non-linear terms, and so the fits are simply those of two differential equations in parallel. Subjects 5 and 10 required the "saturating" effect, Subjects 7 and 9 the "inhibition" term, Subjects 3 and 6 the "potentiation" term; and the fit to Subject 1 was significantly improved by using both the "saturating" and "potentiation" terms.

Table 10.2 gives values for the two gains for each subject calculated for the model's eventual response to steady state values of S of 0, 10 and 15 %. When S=0 the gain is simply that already calculated by the model. The other two values were calculated by substituting values for all the parameters and S into the steady state reduction of the model:

$$Y_1 = G_1 \cdot (1 - I \cdot Y_2) \cdot S$$

$$Y_2 = \frac{G_2}{(1 + \sigma \cdot S)} S$$

and then dividing Y_1 and Y_2 by S to get the gains for that value of S. For example, the G_1 at S=10 for Subject 7 is given by:

$$G_{1(S=10)} = 1.5 (1 - 0.05 \times 0.4 \times 10) / 10$$

$$(\text{as } Y_2 = 0.4 \times 10)$$

This table shows clearly the effect of the non-linear terms on the overall gain for falls in S_{a,O_2} in the range of interest. When considering the response to long term hypoxia, the values of the gains at S=10 and S=15 are arguably more informative than the values calculated directly by the

model program, as when there is non-linearity in the system these latter values will not be representative once S begins to fall.

Table 10.3 shows, for Subjects 1 and 6, the change in each model parameter (starting from the best fit) needed to cause a significant degradation of the fit at the 5% level as calculated using the statistical method of §9.9. This gives a rough guide as to how well each parameter value was defined by the data and model. Several parameters were very well defined, such as G_1 for both subjects, but others (eg τ_2 and I for Subject 1) could take a wide range of values without a significant difference to the fit. The best fit value was not always midway between the two limit values: this was most obvious for τ_2 of Subject 1, where decreasing the time constant by 2 sec caused a significant worsening of the fit, whilst increasing it ten-fold did not. Unfortunately lack of time prevented this analysis from being carried out for all the subjects.

Tables 10.1 - 10.3 overleaf...

Table 10.1 "Best fit" parameter values for the 10 subjects, grouped according to the non-linear terms required in the fit.

Subject	G_1	I	τ_2	G_2	σ
2	0.30	0	70	1.6	0
4	0.05	0	20	0.9	0
8	0.35	0	16	0.65	0
5	0.50	0	70	5.5	0.26
10	0.85	0	30	6.0	0.3
7	1.5	0.05	15	0.4	0
9	1.0	0.25	4	0.4	0
3	0.65	-0.015	8	2.4	0
6	4.5	-0.014	15	5.0	0
1	1.5	-0.001	5	1.5	0.1

G_1 - Gain of equation 1 (l/min/%)

G_2 - Gain of equation 2 (l/min/%)

τ_2 - Time constant of equation 2 (sec)

I - "Inhibition" if positive,
 "Potentiation" if negative ((l/min)⁻¹)

σ - "Saturating effect" (%⁻¹)

Table 10.2 *Gains calculated for steady-state $S=0$, $S=10$, $S=15$. See text for details of calculation. This table shows the effect of the non-linear terms in the model. The "saturating effect" affects G_2 and "inhibition"/"potentiation" affects G_1 , but where these effects are absent the gain does not change with S (indicated by ellipsis ...).*

Sub- ject	G_1			G_2			Total Gain		
	$S=0$	$S=10$	$S=15$	$S=0$	$S=10$	$S=15$	$S=0$	$S=10$	$S=15$
2	...	0.30	1.6	1.9	...
4	...	0.05	0.90	0.95	...
8	...	0.35	0.65	1.0	...
5	...	0.50	...	5.5	1.5	1.1	6.0	2.0	1.6
10	...	0.85	...	6.0	1.5	1.1	6.9	2.4	2.0
7	1.5	1.2	1.1	...	0.4	...	1.9	1.6	1.5
9	1.0	0	-0.5	...	0.8	...	1.8	0.8	0.3
3	0.65	0.88	1.0	...	2.4	...	2.6	3.3	3.4
6	4.5	7.7	9.2	...	5.0	...	9.5	12.7	14.2
1	1.5	1.51	1.51	1.5	0.75	0.6	3.0	2.3	2.1

Table 10.3 *Test of how well each fitted parameter was defined by the data. Each parameter in turn was increased and decreased from its best fit value until the fit was significantly worse at the 5% level. All other parameters were kept at their best fit value. Parameter names and their units are as for Table 10.1.*

Parameter	Lower Value	Best fit Value	Upper Value
-----------	-------------	----------------	-------------

Subject 1:

G ₁	1.2	1.5	1.6
I	-0.003	-0.001	+0.01
τ ₂	13	15	240
G ₂	0.4	1.5	1.7
σ	0.08	0.1	0.4

Subject 6:

G ₁	4.0	4.5	5.4
I	-0.010	-0.014	-0.02
τ ₂	10	15	17
G ₂	4.7	5.0	5.8

10.6 VALIDATION OF THE MODEL

A framework for checking the validity of a model (ie that the model is an adequate representation of the system under consideration) is given by Murray-Smith and Carson (1988). They list a number of criteria that the model should fulfil; these will be discussed in turn with specific reference to the model.

Internal validity: The model itself, without reference to any data or theory, must be mathematically consistent with appropriately chosen algorithms and no internal contradictions. All the algorithms used were checked with test data, and detailed discussion of them may be found in Chapter 9. The basic ideas used in the model (linear differential equations and simple non-linear effects) were straightforward, and similar ideas have been used in many other circumstances.

External validity: This refers to the validity of the model in terms of its purpose, any existing theory and the data. The model should **fit the available data** and **should be compatible with existing theory**. It should **satisfy the objectives** set out for its use and should also have some **potential for "scientific explanation, discovery and hypothesis testing"**. These points will be considered in turn.

The whole of this chapter and much of the last has been concerned with **fitting the data**. The success of this may be seen in the diagrams accompanying these sections. It was unfortunately not possible to give an absolute measure of statistical significance attached to a goodness of fit value (eg as would be obtained by using the chi-squared technique), but the residual sum of squares technique described in §9.9 does allow us to show objectively that the model consistently fits the data better than any of the others tried with a similar or lesser complexity, and fits the data at least as well as any of the more complex ones tried.

Compatibility with existing theory is a problem. There is very little directly relevant quantitative theory, which is why the model was developed the way it was. As described in Chapter 1, §1.2, hypoxia is known to be detected by the carotid body, which responds rapidly to changes in P_{O_2} , and it is known that in the longer term the initial response is altered by other

mechanisms (eg central depression (§1.2.5)). The relationship of \dot{V}_E to S_{a,O_2} has been observed to be linear in rebreathing isocapnic hypoxia studies (Rebuck and Campbell, 1974). The model is not incompatible with these observations, though it is rather more complicated (as would be expected, given that the reason for its development was that the data itself seemed to be at odds with the limited existing theory). The model gives a rapid response to hypoxia, but this response is affected by a second process acting rather more quickly than the central effects observed previously. The structure of the final model is not implausible physiologically: this will be discussed further in the next chapter. Although the model is intrinsically non-linear, in its simplest form, without the "saturating" and "inhibition/potentiation" effects, the relationship between \dot{V}_{Ecalc} and S_{a,O_2} is linear when measured in steady state (ie when S_{a,O_2} remains constant for a period which is long compared with the longer time constant). The non-linear effects, of course, distort this linearity, though not by much when σ or I , or S is small. This deviation from linearity will be discussed more fully in the next chapter.

The **objectives** in developing the model were firstly to find a structure that might shed more light on the discrepancies between hypoxic drive measured using transient hypoxia and that using step hypoxia, and secondly to find a method that could be used for finding a value (or set of values) that summed up a subject's response to short-term hypoxia. The first objective is clearly part of the **scientific potential** validity criterion as well, and will be discussed in the next chapter, but a certain amount of "discovery" and "hypothesis testing" has already taken place during the model's development (the testing of the "inhibition" effect, for example).

The second objective, that of summarising a subject's response to hypoxia, requires further study. The model can be used to obtain 5 (or fewer) parameters for each subject, but this does not mean automatically that the parameters are worthwhile. For this to be the case the parameters must differentiate between the different subjects, ie inter-subject variability must be greater than within-subject variability. To test this, four subjects were studied on ten occasions, three of them once a week on consecutive weeks, the other over a period of 34 weeks. The details of this study are described by Hill (1990) and so are presented only briefly here. The subjects received 4 transient stimuli and then one 15% O_2 step and one 12% O_2 step

(the steps starting with two breaths of 1% O₂ in the same manner as in the rate of change experiments described in Chapter 6). The model was used to fit this data, and the resulting parameter values are shown in Fig 10.9. Analysis of variance (Kruskal-Wallis one-way) showed that inter-subject variability was indeed greater than within-subject variability. Physiological variables frequently vary considerably from day to day (Sahn et al, 1977), so the within-subject variability might well be due to physiological processes, at least in part.

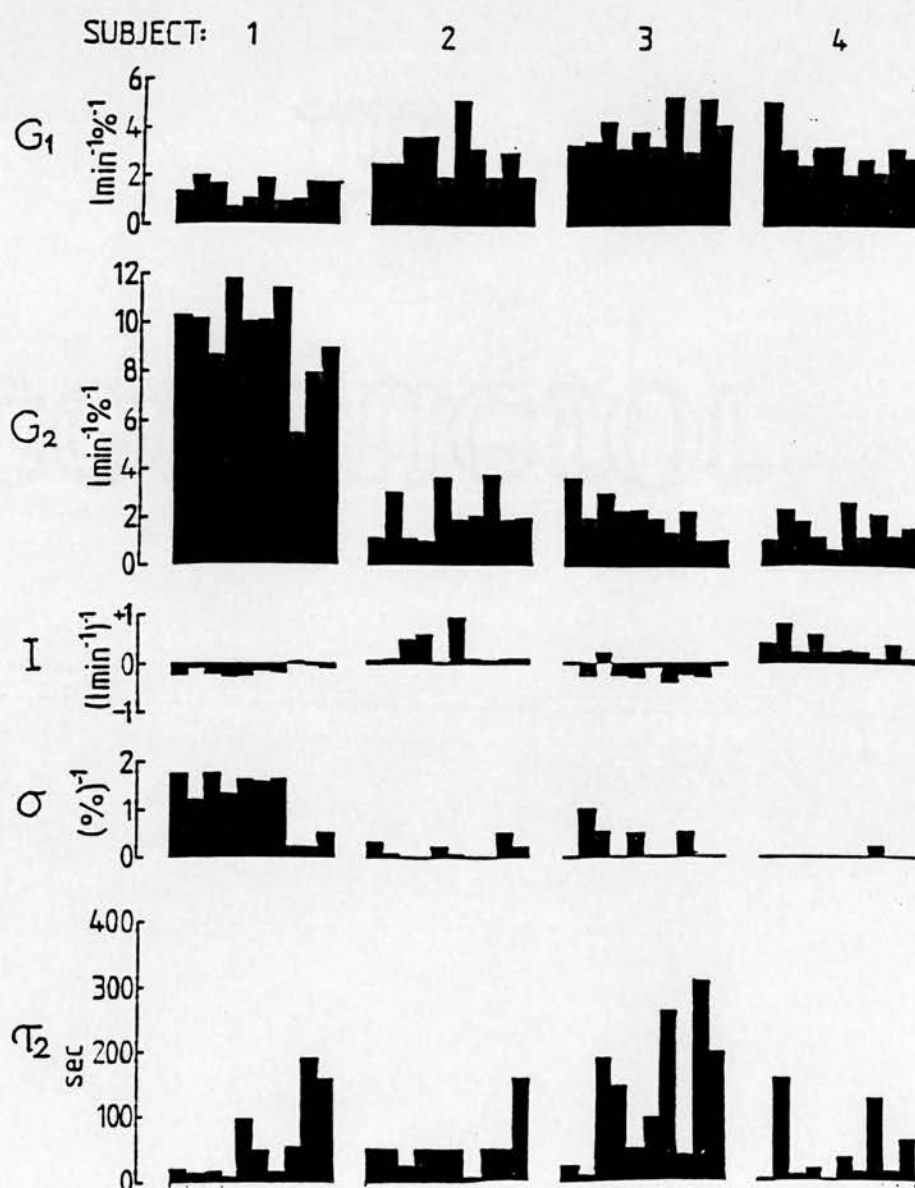


Fig 10.9 Between-day variability of model parameters for each subject on each of ten study days (Hill, 1990).

10.6.1 Validating the model's fit to each subject:

Ideally the parameters obtained for a subject by the model should be tested by using them to calculate curves for a variety of other types of stimulus presented to the subject, and then comparing the calculated responses with those of the subject. Unfortunately the nature of the experiment made this impracticable in anything but a limited way.

There are constraints on the possible shape of the time-course of stimulus for ethical and physiological reasons. Imposing a sudden fall in S_{a,O_2} to much below 80% carries a degree of risk and so must not be done regularly. At the other end of the scale, the natural randomness in \dot{V}_E makes the response to small falls in S_{a,O_2} difficult to measure for most subjects. If the hypoxia lasts much longer than three minutes, effects external to anything contained in the model (such as central depression) are known to affect the response, and the model could not then be expected to fit the data. Practical difficulties (for example the time taken for gases to mix in the lungs, and the awkwardness of controlling P_{ET,CO_2}) limit the fastest rates of change of stimulus. A further difficulty arises due to fatigue in the subjects: adding further stimuli to each study makes the experiment more tiring, and this could affect results; but carrying out a study over more than one session runs the risk of results being affected by changes due to the natural variability of physiological parameters. Within these constraints we have already covered most of the possibilities with our existing experiments. We have used both very brief and longer duration stimuli, extended periods at both low and medium levels of S_{a,O_2} , the fastest possible rate of change of stimulus (on switching to 100% N_2), and a much slower one. The model requires all these types of stimulus in obtaining the set of parameters.

Possible further experiments for use with the model will be discussed in the next chapter. However, a certain amount of checking was possible using the existing data. A subset of the data of each of the subjects, consisting of two steps and three transients, was initially used in finding the parameters for the subject. These parameters were then used as the starting point for a fit to the whole data set. The original reason for doing this was to save on computing time, but it was realised that this also acts as a check on the parameter values. In most cases the parameter values

changed little between fitting the subset and all the data. Although the second pair of steps was usually similar in shape to the first, they were rarely identical as the time course of the fall in S_{a,O_2} was dependent in part on \dot{V}_E at the very beginning of the step (as higher \dot{V}_E meant more rapid mixing of the incoming hypoxic gas with the residual gas in the lungs, leading in turn to a more rapid fall in S_{a,O_2} and a subsequent faster increase in \dot{V}_E). If the parameter values found were highly sensitive to the shape of the input (and therefore worthless for providing a description of the subject), then quite noticeable changes in parameter should sometimes occur between fitting the subsets and the whole sets of data. This is not an ideal test for the parameter values, but it does provide some indication that they are quite robust. [One important exception is in the fit to the data of Subject 10. The second of the two 15% O_2 steps has a considerably greater rise in \dot{V}_E for the fall in S_{a,O_2} than the first, and the model could not fit both together. This observation will be discussed in the following chapter.]

10.6.2 Conclusion:

The validation process just described tests the model against many criteria, and the model cannot be rejected outright on any of them. The main criticism to be raised is in the lack of detailed independent checks on the parameter values obtained for each subject, though there is evidence to show that the parameter values are reasonably robust. Whether the model actually fulfils its initial objectives of explaining the response to hypoxia and providing meaningful information about the subjects of the studies can be judged after the fuller discussion in the next chapter.

CHAPTER 11

DISCUSSION

The previous chapters in this thesis have presented, stage by stage, an argument based on several assumptions that has led to the adoption of one particular mathematical model over all others. This chapter will present a brief summary of the development of the final model, followed by a description of some implications of that model, suggesting possible interpretations with experiments that might refute or support them. Possible uses for the model are discussed briefly before completing the final stages of model validation started in the last chapter.

11.1 SUMMARY OF THE ARGUMENTS USED IN DEVELOPING THE MODEL

1. In the experiments described in Chapter 4, 8 out of 10 subjects showed a greater increase in ventilation (\dot{V}_E) for a given fall in arterial oxygen saturation (S_{a,O_2}) when subjected to 3 minutes of 12% or 15% O_2 gas mixtures than with 3 breaths of pure nitrogen.
2. It was assumed that S_{a,O_2} was a measure of the hypoxic stimulus being reacted to, and that \dot{V}_E was a measure of the response [Chapter 1, §1.2.2, §1.3.1]. In addition, it was assumed that all other ventilatory stimuli (arterial P_{CO_2} , level of exercise, psychological arousal etc.) were constant during the critical parts of the study.
3. By simulating the experiments using an established computer model of the respiratory system and its controllers (Saunders et al, 1980), it was shown that the experimental results could not be accounted for by current knowledge as encapsulated in that model. [Chapter 7]
4. A study presenting the subjects with different initial rates of fall of S_{a,O_2} did not provide evidence of a direct link between rate of fall of S_{a,O_2} and magnitude of response. [Chapter 6]

5. The stimulus-response relationship was shown by several means not to be a simple linear one (ie purely due to a single slow time constant preventing full response being reached during the brief duration of the transient stimuli). Firstly, the system relating \dot{V}_E to S_{a,O_2} must be non-linear as the convolution function was different for a given individual when derived for the transient stimuli compared with that obtained for the step stimuli (in a linear system the convolution function should stay the same) [Chapter 5]; secondly, the response of two subjects was lower during the steps, suggesting a more complicated mechanism at work; and thirdly, a linear mathematical model based on a differential equation with a slow time constant was unable to simulate the observed results. [Chapter 8]

6. Further assumptions were made to allow the nature of the non-linear relationship to be examined mathematically. Most basically, it was assumed that there **could** actually be an underlying mathematical relationship that would describe the effects seen in all the subjects, varying between individuals only in the values of the parameters of the equation, not its form. Following common scientific practice, it was assumed that a simpler relationship was more likely to be of value than a more complicated one giving the same results. To limit the number of possibilities that were to be tried, it was assumed that the terms of the equation would only be ones that mirrored the sorts of broad effects known to occur in nature (saturating effects, inhibition etc.) rather than simply using polynomials of increasing complexity.

7. The aim of the investigation was therefore to find the simplest mathematical relationship that was effective for all the subjects, with a view to using this relationship to help describe the response of each subject, and possibly to gain a better understanding of physiological processes taking place during hypoxia.

8. Without a clear quantitative physiological theory on which to base the model structure, the strategy adopted was to start with the simplest sensible model and increase its complexity in a step-by-step fashion, trying out all possible model structures at each level of complexity and stopping at the simplest structure compatible with the data (which in practical

terms meant stopping at a level where the addition of further complexity did not bring any significant improvement).

9. A simple linear differential equation relating the rise in \dot{V}_E to the fall in S_{a,O_2} was used as the starting point. Other workers (Rebuck & Campbell, 1974) had noted a linear relationship between \dot{V}_E and S_{a,O_2} in steady state experiments, so it seemed reasonable to start by manipulating a basically linear equation rather than to start with something non-linear. Linear differential equations have wide applications in physiology and other branches of science.

10. This linear differential equation was defined by two parameters, the gain and the time constant. It proved impossible to fit adequately the data of a test subject with any combination of values for these two parameters. It was possible, however, to produce a reasonable fit for the steps alone or, with a different gain and time constant, the transients alone, suggesting that this line of approach might be fruitful. [Chapter 8]

11. As the gain appeared to be different between the transients and the steps, various terms were tried in the equation to cause the gain to vary as the S_{a,O_2} fell. The possibilities here were limited by the assumption that the output was **only** dependent on hypoxia as measured by S_{a,O_2} (and, of course, on time). The non-linear terms tried did not make a significant improvement in the fit. [Chapter 8, §8.5]

12. To keep the number of parameters of the equation to a minimum, it was thought preferable to try combinations of two linear differential equations (4 parameters - 2 gains and 2 time constants) rather than adding further non-linear terms to the single equation. There are only a few distinct arrangements of two differential equations, and so all were tried. A dramatic improvement in fit was obtained by using two differential equations "in parallel", with one time constant fast, the other slow. [Chapter 10, §10.2.2]

13. Smaller but significant improvements in fit could be brought about by including a term in the slower differential equation that caused the gain to saturate as S_{a,O_2} fell; and a term inhibiting or potentiating the output of the faster equation allowed an approximate fit to the data of the two subjects showing the greater response to transients than steps, and a

significant improvement in the fits for some of the other subjects. A wide variety of other non-linear terms were tried without any consistent significant improvement in fit. [Chapter 10, §10.2]

14. The output of the slower equation was truncated to zero after S_{a,O_2} reached a value within 2% of baseline. This removed the slow tail-off of the calculated response which conflicted with the rapid cut-off observed (probably caused by the lack of control of P_{a,CO_2} at this point, which therefore brought in a factor outside the simple \dot{V}_E - S_{a,O_2} system being analysed). [Chapter 10, §10.2.4]

15. The resulting model fitted the data consistently better than any of the other models tried, and appeared to be the simplest model using the approach adopted that was able to match both the difference in size of response between the steps and transients, and the general time course of the step response. [Chapter 10]

16. Parameter values were found using the model for all the subjects. A reproducibility study showed that parameter values varied more between subjects than within a subject measured weekly over 10 weeks. [Chapter 10, §10.5]

11.2 THE MODEL IN WORDS

The model is a set of mathematical instructions for converting one set of numbers (representing a changing stimulus) into another set (representing the response). This section attempts to translate these instructions into words in the context of the physiological response to hypoxia. The different elements making up the model are treated in turn.

11.2.1 The two linear differential equations in parallel:

Each of the two differential equations may be thought of as a simple process which detects by how much S_{a,O_2} has fallen from its normal level and reacts to it, with both processes contributing to the response. The response of each is proportional to the size of the stimulus, but one has a much slower response time than the other, ie it responds much more sluggishly, taking many seconds to achieve a full response. When there are no non-linear effects in the model, the two processes are independent

and additive. The gains of the equations then decide the relative importance of each process in the final response, and the sum of the two gains gives the overall hypoxic drive.

The two processes are mathematically simple, but this does not imply that the physiology represented is simple. It is not even possible to deduce that the two mathematical processes each correspond to a separate, single physiological process. It may be the case, for example, that one of the two processes is itself the combined effect of two systems which a physiologist might consider completely separate, but which cause similar changes in response to the stimulus. An example might be two hypoxic detectors in different places, each contributing independently to the response: to a physiologist they might be separate systems, but if they have identical time constants and there is no appreciable time delay between the stimulus reaching each one, then the analysis leading to the model would not be able to distinguish between them. The converse may also be the case. The two model processes might be descriptions of processes occurring within a single physiological entity, perhaps a fairly rapid chemical reaction and a more gradual diffusion occurring at the same site, each affecting the response. Assuming that there **are** actually two physiological systems at work, it is not necessary for each to be detecting S_{a,O_2} directly. The model suggests that each process receives the same stimulus, but it is possible that there is only one place where oxygen levels in the blood are sampled directly, which provides a neural output which is fed to the two systems simultaneously.

The mathematical analysis may be unable to distinguish the number of physiological sites where detection of hypoxia takes place, but the fact that there are **two** types of process occurring is an important step forward. It shows that response to hypoxia is not just a very simple stimulus-response effect, but nor is it so complicated (or random) as to be beyond experimental elucidation. It suggests that studies should be directed towards isolating the two processes, and invasive studies that provide information about a single simple hypoxic centre cannot conclude they have a complete description of the response to hypoxia until at least two processes have been identified. Further studies on this will be suggested in a later section of this chapter (§11.3).

11.2.2 The Saturating Effect:

In many of the subjects the gain of the slower-acting differential equation decreased as the stimulus increased. In terms of the two processes described above, the saturating effect means that the slower process, quite independently of the faster one, responds proportionally less to a large fall in S_{a,O_2} than to a small one.

11.2.3 Inhibition:

The inhibition term was put into the equations in an attempt to make the model output match the data from two of the subjects in which the response to transient hypoxia was proportionally greater than that to step hypoxia. This term causes the gain of the faster process to decrease as the response of the slower process increases. As the S_{a,O_2} falls at the start of hypoxia, the output of the fast process begins to rise, as does that of the slower process, though more gradually. If the S_{a,O_2} falls rapidly, the faster process dominates the response, and if the S_{a,O_2} returns quickly to baseline, as in a transient, the slower process has insufficient time to inhibit the response. If, however, the S_{a,O_2} falls more slowly and stays low, as in a step, the output of the slower process increases and cuts down the response of the faster process. The overall output is therefore less than it would have been for the same fall in S_{a,O_2} when inhibition did not take place.

As was noted in §10.3.1, the data for the subjects requiring the inhibitory term was not good enough to be able to differentiate between possible mathematical descriptions of exactly how the inhibition takes place. The method used was chosen for its simplicity: less noisy data may show that an alternative is better.

For inhibition to work in the way described puts some restraints on the parameters of the model. If the time constant of the slower process is very much greater than that of the faster process then there will be a peak in response at the beginning of a step before the inhibition of the fast process begins. This peak will be more pronounced as the rate of increase of stimulus increases (indeed it is the cause of the large response to the transient stimulus where the fall in S_{a,O_2} is very rapid). The data is too

noisy to confirm or refute the possibility of this peak existing for the two subjects. The study described in Chapter 6 (where the rate of fall of S_{a,O_2} at the start of a step is increased by starting with two breaths of 1% O_2) would show whether this peak occurs, but unfortunately these two subjects were not available for testing.

A further restraint is on the relative sizes of the gains of the two processes and the inhibition parameter. These need to balance so that the reduction in output of the fast process is greater than the amount added to the response by the slow process, but yet is not so great as to end up with a negative overall response.

Conclusions about the inhibitory term can only be tentative. Arguments in §10.3.1 suggested that it is likely that some form of inhibition was occurring in the two subjects, and the model describes one way that this might work. It is quite possible, however, that something other than the second process (but with a similar time constant) is inhibiting the first. Only data of a very good signal-to-noise ratio would be able to show this, and only if the inhibitory and the slow process had sufficiently different time constants.

11.2.4 Potentiation:

In most of the subjects the difference in response to steps and transients resulted in the model from the slower differential equation having a combination of a large gain and a long time constant. Two subjects, however, had an extremely large response to steps which increased rapidly right from the start of a step. For these two subjects the time constant of the slower process was too fast to allow a sufficient difference in overall response: increasing the gain increased the response during transients as well as steps, whilst increasing the time constant prevented the fast initial rise in response at the beginning of a step.

The potentiation term was added to the model to help fit this type of response. As the slow process increases in response, the potentiation causes an increase in gain in the fast process. During a transient, when the slow process makes little contribution, the potentiation has little effect; but during a step the slow process not only adds to the total response itself, but

also causes the fast process to become more sensitive to the fall in S_{a,O_2} and so to increase the overall response even more.

The potentiation term in the model is treated exactly as negative inhibition. This is convenient for parameter fitting, but it does not necessarily mean that the inhibition and potentiation are alternate facets of the same physiological process. The discussion on inhibition (above) stated that there was the possibility that the inhibition results from a third process with a similar time constant to the slow process. The same argument can be applied to potentiation.

11.3 PHYSIOLOGICAL INTERPRETATION OF THE MODEL

The model as just described can be used directly to characterise a subject's response to short hypoxic stimuli, but one of the aims of modelling, possibly the most interesting from the physiologist's point of view, is to relate the theoretical "processes" of the last section to known (or possibly even new) physiological mechanisms.

The ideas of the last section were all drawn directly from the model and therefore ultimately from the data. In this section the ideas are purely speculative and can only be accepted or rejected on the basis of future studies designed to test them. If a hypothesis is contradicted by the model it can be rejected, but any number of hypotheses may agree with the model. The advantage of using the model is that it provides a basis for clear thought by summarising the characteristics required of a complete hypothesis.

11.3.1 The fast process:

Animal studies and studies of human subjects whose carotid bodies have been resected (Wade et al, 1970; Lugliani et al, 1971) have established that the carotid bodies are of fundamental importance in detecting and responding to hypoxia. Ventilation increases rapidly as hypoxic blood reaches them, the response time being of the order of a few seconds (Swanson and Bellville, 1975; Bellville et al, 1979; Gardner, 1980). These observations must find a place in the model, and the most sensible candidate is the fast process. This process displays a quick reaction to the fall in S_{a,O_2} which is linear when the hypoxia is transient (when there is

insufficient time for any inhibition or potentiation to become established). Certainly the slow process generally has too long a time constant to be a model of the chemoreceptors, but, as was argued earlier, it may be possible that the fast component describes more than just the carotid body-mediated response, and contains other responses that have a similar time constant.

The simplest hypothesis would therefore be that the fast process in the model is solely a reflection of the activity of the peripheral chemoreceptors.

This hypothesis can be tested in a variety of ways, some easier than others. The aim must be to carry out an experiment where the response of the peripheral chemoreceptors is affected in a known way, the ideal being to remove it from the system altogether. If the hypothesis is correct, the gain of the fast process should vary in exactly the same way as the known effect on the receptors, ie both should increase or decrease proportionately. When the response of the chemoreceptors is removed from the system entirely, the gain of the fast process should be zero. The effect of this on the slow component depends on the mechanism behind it and will be discussed below.

The response of the peripheral chemoreceptors can be manipulated by a number of methods. If the model were shown to be applicable to the hypoxic response of anaesthetised cats (or other animals), a direct comparison could be made between the hypoxic response before and after resection of the carotid bodies. It is unlikely, however, that the model would actually be valid in such circumstances given that it was developed for conscious exercising humans.

A more practical solution is to alter the response of the chemoreceptors chemically, preferably during an exercise study similar to the one on which the model was based. This might be done pharmacologically by using a drug such as Almitrine which specifically increases carotid body activity (Laubie and Schmitt, 1980), comparing studies with and without the drug. Simpler still is to use one of the substances that has greatest effect on the carotid body: oxygen.

Transient hyperoxia: Experiments (Dejours, 1964; Stockley, 1978) have shown that breathing 100% O₂ stops neural output from the carotid body. If two breaths of 100% O₂ are given to the subject towards the end of a hypoxic step (after sufficient data for that step has been collected to allow the program to find the model parameters), then if the hypothesis is correct, \dot{V}_E should fall transiently, probably with a time constant similar to that of the fast process (ie faster than can be measured). At the lowest point \dot{V}_E should be exactly that expected of the slow process alone, ie approximately equal to the gain (G_2) times the fall in S_{a,O_2} (S) plus the baseline \dot{V}_E . \dot{V}_E would be expected to rise towards its previous level once the carotid body recovered its normal level of sensitivity. Because of its slow time constant, the slower process should not be noticeably affected by the change in O₂, just as it is not affected during transient hypoxia. This experiment is, in effect, a mirror image of the standard transient hypoxic technique described earlier in this thesis.

The size of the fall in \dot{V}_E may not be as expected. If there is no fall in response, then it is unlikely that the fast process is a representation of the action of the carotid bodies. If the fall in \dot{V}_E is not as large as expected, then the fast component is likely to involve a mechanism in addition to the carotid body response (although an alternative explanation might be that the hyperoxia was not sustained enough to completely cut out carotid body activity). Less likely, on the other hand, is the result that the fall in \dot{V}_E is greater than expected, which would indicate that the slow process is dependent on the carotid body in quite a complicated way (as this implies a faster time constant for the slow process during hyperoxia than hypoxia - otherwise the effect would not be noticeable).

A fall in \dot{V}_E to the predicted level in the above experiment would not in itself prove the hypothesis; but since the carotid body must appear somewhere in the model and its time constant of response is fairly short (so that it cannot be hidden in the slow process) the conclusion would seem reasonable. A result at odds with the hypothesis would, however, be interesting.

It may be argued that, strictly, a positive result as described does not preclude the possibility of a second chemoreceptor with properties (time constant, time lag and response to 100% O₂) similar to the carotid body's.

The response from such a receptor should, however, have shown up clearly in hypoxia studies on carotid body resected subjects. Swanson et al (1978) showed a very small (2 l/min) increase in \dot{V}_E in carotid body resected subjects with step hypoxia, but only during hypercapnia. This may have been due to the aortic chemoreceptors, but the effect is too small to be of importance here.

11.3.2 The slow process:

This is perhaps more interesting than the fast process as it does not have a single, obvious explanation in current theories of hypoxic response. The properties of the slow process that a theory would have to incorporate are: a slow time constant that can vary considerably between individuals (8 to 70 seconds in the subjects tested initially); a gain generally larger than that of the fast process (which again can vary markedly); a possible upper limit to its contribution to the overall response (the saturating effect); and the possibility of its response inhibiting or potentiating that of the fast process.

There are many possible mechanisms that might explain the slow process. Most involve diffusion in some form or another: in nature, slow time constants are often associated with diffusion. This makes further study of the process difficult without some form of invasive technique for establishing what is diffusing where.

Possible mechanisms can, for convenience, be divided into five groups according to how directly the slow process is related to P_{O_2} : those that are due directly to the reduced P_{a,O_2} , those that follow directly from the detection of the reduced P_{a,O_2} , those that are secondary to the reduction in P_{a,O_2} or secondary to the detection of the reduction, and those that are independent of the hypoxia (ie are artifactual). Some examples of each follow.

1. Mechanisms due directly to the fall in P_{a,O_2} :

This type of mechanism implies a detector of reduced P_{a,O_2} which, to give the long time constant, either responds gradually to the fall in P_{a,O_2} or is responding to an environment where P_{a,O_2} is gradually falling (an example of the latter would be when hypoxic blood diffuses only slowly into the receptor). As carotid body resected subjects do not display a

response to hypoxia large enough to account for the slow component, this cannot be due to a second hypoxia detector. This results in the hypothesis that the carotid body response itself involves two components. The carotid body is very well perfused (Daly et al, 1954), so the slow component is unlikely to be due to the slow diffusion of blood or of oxygen between tissue and blood. It is possible that there may be a slow release of a chemical in the carotid body in response to the hypoxia which increases its sensitivity; for example, it is known that the neurotransmitter dopamine is released from the carotid body during acute hypoxia (Fidone et al, 1980), and although this cannot be the cause of the slow component (dopamine is inhibitory), it suggests that this sort of mechanism might be possible. Verification of this could only be carried out invasively in animal studies.

2. Mechanisms due directly to the detection of the fall in P_{a,O_2} :

Until the inhibition term was added to the model, two equally good arrangements of the two differential equations were possible, one of which involved the output of the faster equation being the input to the slower one - in other words, the slow process might be a reaction to the response of the carotid body to the fall in P_{a,O_2} . An example of this might be a neurological mechanism where signals from the carotid body not only cause an immediate increase in respiratory activity, but are also detected centrally where, if hypoxia is prolonged, a separate drive to respiration is set off. Millhorn et al (1980) found a long term increase in neural activity in cats following transient stimulation of the carotid body. It is possible that a similar mechanism may occur in man.

It is important to note that if it is necessary to incorporate inhibition/potential in such a hypothesis, this would be inhibition or potentiation of the pathway linking carotid body and increased ventilation, and not directly of the output of the carotid body as this could lead to an unstable or oscillating system.

It is not easy to design an experiment to test such a hypothesis. There is no reason to suppose that the time constant of the decay of the second process following the end of hypoxia should be the same as the time constant at onset, so if a study where hypocapnia at the end of hypoxia was carefully controlled showed a changing time constant, this might support this hypothesis.

3. Mechanisms secondary to the fall in P_{a,O_2} :

The fall in P_{a,O_2} is likely to cause other chemical changes in the body, and these might sensitise the carotid body or affect other respiratory mechanisms. For example a decrease in pH in the blood, due perhaps to the muscles working in a hypoxic environment, would stimulate the respiratory centre in the medulla as if the blood were hypercapnic. The slow time constant could be due here either to a slow chemical reaction or to the slow diffusion of the chemical to the respiratory centre (the central response to hypercapnia displays a similarly slow time constant).

One argument against this hypothesis is that if the chemical change is due to reactions in the muscles, there ought to be a noticeable delay (dependent on cardiac output, but certainly greater than 10 seconds) after the onset of hypoxia before the slow component starts to take effect (due to the time for the hypoxic blood to reach the muscle and the chemical to circulate back to the chemoreceptor), yet this is not found in the data. Testing the more general hypothesis requires the sampling of arterial blood at intervals during hypoxia. This is awkward and carries some risk, but is possible. The blood should be tested for changes in pH, and any chemical thought to be a potential respiratory stimulant. A build up of such a chemical at a rate consistent with the time constant of the slow process would suggest potentiation of the carotid body, while a faster build up would suggest that the mechanism involves diffusion of the chemical to the central receptor.

An alternative to the chemical signal is the possibility of increased neural signals being received from the exercising muscles due to the hypoxia, or even from the respiratory muscles (Gandevia and Macefield, 1989). This mechanism would not be subject to the same length of delay as the chemical mechanism just described, but there should still be a delay of a few seconds before hypoxia begins to take effect in the muscles, and again this is not seen in the data. Testing this hypothesis requires the selective denervation of animals. Before such experiments would be worthwhile, it would have to be clearly established that the two component response to hypoxia is valid for the animal being tested.

It is possible that a physiological rather than chemical or neural change happens because of the hypoxia, altering the response of a chemoreceptor.

For example, hypoxia could affect the diameter of blood vessels which could in turn alter the central response to P_{a,CO_2} due to changes in CO_2 elimination by the brain (Nishimura et al, 1987). Even if such changes did take place in such a way as to increase ventilation, it seems unlikely that the increase would be as large as is seen in the response of some of the subjects.

4. Mechanisms secondary to the detection of the fall in P_{a,O_2} :

Hypoxia, once detected, leads to an increase in rate or depth of respiration, and the subject may notice this consciously or unconsciously, leading to a psychological effect on ventilation. The subject can deduce that the stimulus has been applied, and arousal or stress might result in a further increase in ventilation. An increase in ventilation is frequently associated with stress. Increases of catecholamines in the blood during hypoxia might support this being at least a factor in the slow process.

Hypoxia can also affect cardiac output, and therefore the blood flow at the chemoreceptors, which might cause a change in their response. If, for example, the carotid bodies react to the rate of change of P_{a,CO_2} during oscillations (caused by the respiratory cycle (Cross et al, 1984)), then continuously increasing cardiac output will cause an increase in \dot{V}_E . This is because the rate of change of P_{a,CO_2} at the carotid body will increase as the time difference between trough and peak diminishes due to the faster blood flow. It seems unlikely that this would be a large enough effect to explain the slow process, and it is difficult to test whether it actually is a cause as a gradual increase in cardiac output during the course of step hypoxia could be a response to the slow process rather than a cause of it. If the cardiac output were to continue rising once the slow process had reached its maximum, this would count against this hypothesis.

5. Mechanisms independent of the fall in P_{a,O_2} :

It is possible that the slow process is due to the experimental method rather than changes due directly or indirectly to hypoxia. One possible cause might be the incorrect control of P_{a,CO_2} . This will be discussed in some depth later and in Appendix A.

11.3.3 The saturating effect, inhibition and potentiation:

Interpretation of the non-linear terms depends heavily on the mechanism causing the slow component. For example, if the slow process were due to chemical changes then potentiation might result from neurological sensitising of the carotid body following detection of the chemical, or by direct stimulation by the chemical itself. If, on the other hand, the slow process were a psychological effect, the potentiation might be explicable in terms of that effect (it may be that the psychological component is not a simple linear process and the "potentiation" is actually just a demonstration of this rather than a strict potentiation of the fast process). Until an acceptable hypothesis explaining the slow process is agreed, it is not profitable to try interpreting the inhibition and potentiation terms in isolation. Both potentiation and inhibition due to central mechanisms have been demonstrated in cats (Tenney and On, 1977)

Some more general interpretations might be put on the saturating effect. This could be due to respiratory fatigue, or could be a demonstration of the optimising model of Poon (1983) in which \dot{V}_E is set to a level that takes into account both the degree of hypoxia and the increased work involved in breathing. It would be interesting to find maximal \dot{V}_E for the subjects and see if this relates to σ . The saturating effect could be a manifestation of diffusion somewhere approaching a maximum level as saturation is reached. Another possibility is that the second process is simply reacting more linearly to the fall in P_{a,O_2} than to S_{a,O_2} , and that the curved relationship with S_{a,O_2} is simply due to the shape of the dissociation curve relating P_{a,O_2} and S_{a,O_2} .

11.4 THE CONTROL OF ARTERIAL P_{CO_2}

CO_2 is an important respiratory stimulant, acting both centrally and at the carotid bodies, where it also increases the sensitivity to hypoxia (see §1.2.2). Hypocapnia, a reduction in P_{a,CO_2} , tends to reduce the hypoxic response. It is therefore essential during hypoxia studies that mean P_{a,CO_2} is maintained at a constant level, and it was an assumption during the development of the model that this was the case in our studies.

An increase in \dot{V}_E during hypoxia will cause an increase in the amount of CO_2 expired. To prevent hypocapnia, CO_2 was added to the inspired gas at a concentration sufficient to keep $P_{\text{ET},\text{CO}_2}$ constant. The use of $P_{\text{ET},\text{CO}_2}$ as an index of changes in mean P_{a,CO_2} is widespread, even though it is acknowledged that on exercise $P_{\text{ET},\text{CO}_2}$ and mean P_{a,CO_2} are not equal. The assumption being made is that the difference between $P_{\text{ET},\text{CO}_2}$ and mean P_{a,CO_2} remains approximately constant even when \dot{V}_E changes. If this is **not** the case then keeping $P_{\text{ET},\text{CO}_2}$ constant is not appropriate, and doing so will alter mean P_{a,CO_2} and thus the overall stimulus. If the difference between $P_{\text{ET},\text{CO}_2}$ and mean P_{a,CO_2} were to decrease with either increasing \dot{V}_E or increasing inspired CO_2 , then keeping $P_{\text{ET},\text{CO}_2}$ constant would result in an increase in P_{a,CO_2} , thereby increasing the stimulus. At worst, the increased stimulus might result in a further change to the $P_{\text{ET},\text{CO}_2}$ - P_{a,CO_2} difference ($[\text{ET-a}]$), and thus eventually to a further compensating increase in stimulus. A small effect could therefore be magnified over the course of several breaths.

It is clearly important to establish under what circumstances the $[\text{ET-a}]$ difference changes, and whether or not this might be affecting our results. The arguments for this are necessarily detailed and so have been consigned to Appendix A, with only a summary given here.

Ideally, a check on the effectiveness of the control of P_{a,CO_2} would involve sampling arterial blood during hypoxia and analysing it for P_{a,CO_2} . This was not possible at the time of the studies. The appendix describes instead an indirect method which compared changes in breath-by-breath \dot{V}_{CO_2} with what was predicted by theory. When applied to a study involving seven 12% O_2 steps, carried out exactly as in the main study described in Chapter 4 with the control of $P_{\text{ET},\text{CO}_2}$ being done by the same experimenter as before, it was found that breath-by-breath \dot{V}_{CO_2} decreased during the periods in which CO_2 was added to the inspired gas, indicating that the level of inspired CO_2 was inappropriately high, even though $P_{\text{ET},\text{CO}_2}$ was kept at its normoxic level.

Reasons for this are explored in the appendix. There, a simple model of CO_2 transfer between lung and blood (based on the work of Dubois et al (1952)) is described and used to simulate different patterns of breathing. With this model, the difference between $P_{\text{ET},\text{CO}_2}$ and mean P_{a,CO_2}

diminishes with increased inspired CO_2 and with increased breathing frequency, increases with increased tidal volume, and is affected by the shape of the flow profile during breathing. If mixing in the alveoli is not instantaneous, as is usually assumed in models, then this could also be affecting the difference.

The appendix concludes that keeping $P_{\text{ET},\text{CO}_2}$ constant during hypoxia is likely to introduce a systematic change in stimulus. Any study assuming that $P_{\text{ET},\text{CO}_2}$ is a measure of mean P_{a,CO_2} must be treated with caution.

Whatever method is chosen for keeping mean P_{a,CO_2} constant, it is an implicit assumption that this is sufficient for removing any effect CO_2 might be having on \dot{V}_{E} . This may not be the case. Certainly keeping mean P_{a,CO_2} constant should keep the central CO_2 response approximately constant (as long as blood flow in the brain does not change sufficiently to alter the P_{CO_2} of blood reaching the central receptor). It is possible, however, that the carotid bodies may be responding, at least in part, to the waveform of P_{a,CO_2} caused by the changing P_{CO_2} in the lung during the respiratory cycle (§1.2.3). Adding CO_2 to the inspired gas mixture affects the shape of the P_{a,CO_2} oscillations, and so might be affecting \dot{V}_{E} indirectly. It is therefore hard ever to be certain that the stimulus during a study is purely due to hypoxia.

11.4.1 Applying these ideas to our results

It is hard to assess retrospectively whether the control of P_{a,CO_2} has had much influence on the results used for developing the model. Inadequate control of P_{a,CO_2} might explain the response to the second 15% O_2 step in Subject 10, which was considerably greater than during the first 15% O_2 step for a given fall in S_{a,O_2} , the difference being so great that the model was unable to fit this step adequately (see Figs 4.2j and 10.8j). The CO_2 added to the hypoxic mixture could account for the very much increased response of Subject 3 to step hypoxia than transient hypoxia, a difference which manifests itself within two breaths of the start of the response, although at that stage transient and step hypoxia would not be differentiable on the basis of duration. If the CO_2 added to the first few breaths of a step does affect the amount of CO_2 required for the rest of the step to keep $P_{\text{ET},\text{CO}_2}$ constant, this might explain the marked day-to-day variability within a subject of some of the parameter values.

An interesting experiment might be to find the time constants of the subjects' responses to an iso-oxic step increase in inspired CO_2 : this might relate to the time constant of the second process of the model.

11.5 USE OF THE MODEL

There were two original aims in developing the model. One was to help gain a greater understanding of the physiology underlying the response to hypoxia: this has been discussed in the preceding sections. The other aim was to provide a way of summarising numerically the response of a subject to short-term hypoxia so that subjects might be compared with one another, and changes in the response of one subject with time or in different circumstances might be measured. The success of this was discussed in §10.5. Assuming that the potential problems in the control of Pa_{CO_2} described in the last section are not found to be of importance, this section describes briefly the use to which the model could be put.

One study already complete (Hill, PhD Thesis, 19⁹~~8~~0) was an investigation of familial aspects of the hypoxic ventilatory response. This was prompted by observations (Flenley et al, 1970; Bradley et al 1979) that patients with the "blue and bloated" form of chronic obstructive pulmonary disease (COPD) had a low hypoxic ventilatory drive. These patients have low Pa_{O_2} , high Pa_{CO_2} , cor pulmonale and pulmonary hypertension, which distinguish them from patients with the "pink puffer" form of COPD who have relatively normal arterial gas tensions. Hill examined the hypoxic response of first generation offspring of sufferers from COPD, looking for a correlation between values taken by the model parameters for the offspring and the Pa_{O_2} of their parents.

The model might also be used to find out more about hypoxic drive, testing factors that could affect it such as fitness, age, emotional state, food, drugs and so on. If a low hypoxic drive was found to render an individual liable to illness, the model could potentially be used with the results of intervention trials of any drug thought to have a beneficial effect.

On a more theoretical tack, the model might be used for improving existing models of respiration. A model such as that of Saunders et al (Chapter 7) could benefit from a more detailed hypoxic controller. In

fitting our model into the larger one, the only adaptation necessary would be to replace the time constant of the faster process (which could not be found accurately due to the time constant of the ear oximeter) with a value similar to that of the oximeter (2 or 3 seconds, for example). The interaction between hypoxia and hypercapnia would have to be found by trial and error, using comparisons between the output of the model and the \dot{V}_E of subjects with similar parameter values to those in the model exposed to various levels of hypoxia and hypercapnia.

11.6 IN CONCLUSION - THE FINAL VALIDATION OF THE MODEL

Model validation was discussed in Chapter 10. Validation involves fulfilling a wide range of conditions from checking the algorithms to less obvious ones such as fitness for purpose. In that discussion the model was shown to stand up to most of the required conditions, but one condition could not at that time be tested: whether the model fulfilled its two initial aims of helping to explain the physiology behind the response to short-term hypoxia and providing meaningful figures that sum up a subject's hypoxic response. The section on interpreting the model in this chapter (§11.3) shows that the model, at the very least, provides a framework for thinking about and discussing possible physiological mechanisms underlying the response to hypoxia. It is possible that the results of experiments suggested in that section will shed considerable light on some of the more complex parts of the hypoxic response, though whether they do or not, the very fact that these suggestions can be made so specifically shows the benefit of developing the model.

The success of the second aim - of providing meaningful figures summing up a subject's hypoxic response - can only really be judged once the physiology underlying the model is understood more fully. However, if the model can provide a reproducible set of parameters for a subject that can be compared with those of other subjects (and which, unlike earlier measurements of hypoxic drive, do not depend on the test stimulus), then the model may be considered useful. At the time of writing the model is being used regularly in just this way.

The route to a valid model was long and complex, but if it now provides better understanding of the physiology of hypoxia and diseases associated with it, the model can be considered a success.

APPENDIX A

THE CONTROL OF P_{CO_2} DURING HYPOXIA

A.1 INTRODUCTION

\dot{V}_E increases during hypoxia, increasing the amount of CO_2 expired and thus leading to a decrease in mean $P_{\text{a,CO}_2}$. This reduces the response of the carotid bodies to the hypoxia, and the peripheral and central CO_2 responses will also be diminished. If the purpose of an experiment is to measure purely the response to the hypoxic stimulus, it is clearly important that action is taken to prevent the fall in $P_{\text{a,CO}_2}$.

A frequently used method for this is to add CO_2 to the inspired gas mixture during hypoxia in such quantity as to keep end tidal P_{CO_2} ($P_{\text{ET,CO}_2}$) at its pre-stimulus level. The implicit assumption made in using $P_{\text{ET,CO}_2}$ as the controlled variable is that the [ET-a] difference will not vary during the course of an experiment. The $P_{\text{ET,CO}_2}$ method is relatively simple, uses straightforward equipment (a CO_2 detector, a signal display device and a method of controlling CO_2 flow), and can be automated if required (Robbins, 1984).

At rest, $P_{\text{ET,CO}_2}$ and mean $P_{\text{a,CO}_2}$ are similar in value, although the difference between them is known to increase with level of exercise (Jones et al, 1966; Whipp and Wasserman, 1969). Unfortunately, the difference between $P_{\text{ET,CO}_2}$ and mean $P_{\text{a,CO}_2}$ varies with breathing pattern. Jones et al (1979) measured mean $P_{\text{a,CO}_2}$ by sampling arterial blood in 5 subjects, and compared it with $P_{\text{ET,CO}_2}$ at three different breathing frequencies each at two levels of exercise. They found that the [ET-a] difference increased with V_T . This shows that keeping $P_{\text{ET,CO}_2}$ constant does not imply that mean $P_{\text{a,CO}_2}$ is kept constant, especially if V_T changes.

The data on which the development of the model of Chapter 10 was based might therefore include a component in the response which was affected by a change in mean $P_{\text{a,CO}_2}$. The model development, however, assumed that hypoxia was the only stimulus.

This appendix describes an experimental and a theoretical investigation of the adequacy of keeping P_{ET,CO_2} constant as a method for controlling mean P_{a,CO_2} .

A.2 BREATH-BY-BREATH \dot{V}_{CO_2}

A.2.1 Introduction

At the time of our studies it was not possible to check mean P_{a,CO_2} directly using arterial catheterisation. An independent assessment of constancy of mean P_{a,CO_2} could, however, be gained indirectly by looking for changes in breath-by-breath \dot{V}_{CO_2} .

Once steady state gas exchange has been achieved for a given level of moderate exercise, \dot{V}_{CO_2} should not change by much over the course of a few minutes, as the work done by the muscles, and therefore the CO_2 they output, should remain constant. The volume of CO_2 leaving via the lungs in a given time should be equal to the amount being created in the muscles and elsewhere during that time. If this is not so, the amount of CO_2 stored in the body (eg in the lung or dissolved in the blood) must be changing.

The relationship between mean \dot{V}_{CO_2} and mean C_{a,CO_2} (the concentration of arterial CO_2 , a value directly related to P_{a,CO_2}) over a time period T (long enough for fluctuations in the volume of CO_2 in the lung to be averaged out) can be obtained simply as follows:

$$\text{mean } \dot{V}_{CO_2} = \frac{1}{T} \int_0^T (\dot{V}_{CO_2\text{out}} - \dot{V}_{CO_2\text{in}}) dt$$

(where $\dot{V}_{CO_2\text{in/out}}$ are the flow of CO_2 into and out of the lung via the mouth).

Arterial CO_2 concentration (C_{a,CO_2}) at a given instant =

$$\frac{\text{CO}_2 \text{ flow from venous blood} - \text{CO}_2 \text{ flow into lungs}}{\text{blood flow}}$$

$$\text{so: mean } C_{a,CO_2} = \frac{1}{T} \int_0^T \frac{(\dot{V}_{CO_2\text{blood}} - \dot{V}_{CO_2\text{lungs}})}{Q} dt$$

(where $\dot{V}_{\text{CO}_2\text{blood/lungs}}$ are the flows of CO_2 arriving in the venous blood and leaving the blood into the lungs, and \dot{Q} the blood flow to the lungs).

If $\dot{V}_{\text{CO}_2\text{blood}}$ and \dot{Q} are constant over period T , then

$$\text{mean } C_{a,\text{CO}_2} = \frac{\dot{V}_{\text{CO}_2\text{blood}}}{\dot{Q}} - \frac{1}{T \cdot \dot{Q}} \int_0^T \dot{V}_{\text{CO}_2\text{lungs}} dt$$

If T is chosen so that the volume of CO_2 in the lungs is the same at the end as at the start (so T must include a whole number of breaths) then:

$$\int_0^T (\dot{V}_{\text{CO}_2\text{out}} - \dot{V}_{\text{CO}_2\text{in}}) dt = \int_0^T \dot{V}_{\text{CO}_2\text{lungs}} dt$$

so:

$$\text{mean } C_{a,\text{CO}_2} = \text{mean } C_{v,\text{CO}_2} - \frac{1}{\dot{Q}} \text{mean } \dot{V}_{\text{CO}_2}$$

Rearranging gives:

$$\dot{V}_{\text{CO}_2} = \dot{Q} (C_{v,\text{CO}_2} - C_{a,\text{CO}_2}) \quad \dots \text{Equation 1.}$$

(where all the values are averages over the interval T).

Equation 1 is equivalent to that used in the Fick method for calculating cardiac output (see, for example, Jones et al (1975)). The important reason for going through the derivation in detail is that it shows the circumstances under which Equation 1 is valid. In particular, it is valid even when CO_2 is being added to the inspired gas. The equation, however, is not exact when cardiac output alters during time T .

If \dot{V}_{CO_2} is calculated breath-by-breath, it would be expected to fluctuate slightly from one breath to the next due to variations in the breathing pattern, but the underlying trend should follow that described in the equation, as long as cardiac output varies slowly relative to the length of a breath. If mean P_{a,CO_2} is not to change, mean breath-by-breath \dot{V}_{CO_2}

should remain constant while cardiac output and level of exercise are constant.

A.2.2 Methods

The data acquisition program was adapted to calculate breath-by-breath \dot{V}_{CO_2} (\dot{V}_{CO_2br}) as

$$\dot{V}_{CO_2br} = \frac{1}{T_{TOT}} \sum^{breath} F_{CO_2} \times flow$$

where F_{CO_2} is the gas fraction of CO_2 measured at the mouth using the mass spectrometer, T_{TOT} is the length of the breath, and flow is the pneumotachograph output multiplied by a calibration factor (as described in the calculation for tidal volume, §2.2.1).

Seven steps of approximately 50 breaths of 12% O_2 were administered to an exercising subject exactly as described in Chapter 4, except that on three of the steps about 1% CO_2 was added to the inspired gas with the initial hypoxic breath. The reason for this was to see if an initial excess of CO_2 resulted in a greater level of CO_2 being required throughout the step in order to keep P_{ET,CO_2} constant. This would be the case if a change in breathing pattern resulting from the initial added CO_2 decreased the difference between P_{ET,CO_2} and mean P_{a,CO_2} , leading to a positive feedback loop where keeping P_{ET,CO_2} constant meant that P_{a,CO_2} was increased, leading to an increased stimulus and thus a further decrease in the [ET-a] difference, requiring yet more CO_2 to keep P_{ET,CO_2} constant.

CO_2 was added to the inspired gas mixture during each step to keep P_{ET,CO_2} constant, just as in the earlier experiments. The same experimenter as before controlled the flow, and was unaware of the objectives of this study.

A.2.2 Results

There was no discernable difference after the fifth breath between the three steps with initially increased CO_2 and the others.

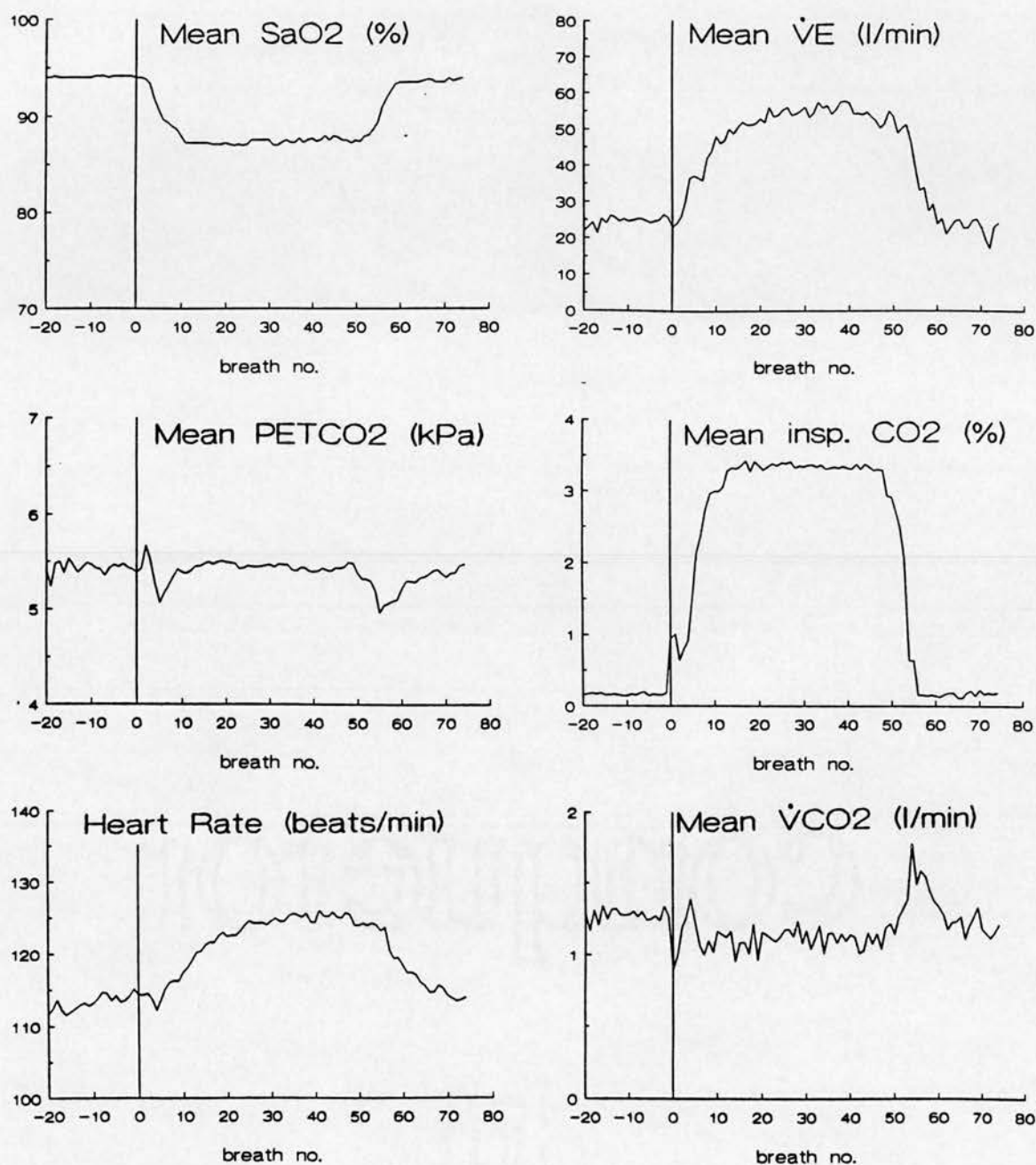


Fig A.1 Plots of curves derived from the study described in A.2.1.1. Each curve is the mean of seven superimposed steps aligned according to the start of the hypoxia (breath 0).

Mean curves for the seven steps were found for \dot{V}_E , S_{a,O_2} , P_{ET,CO_2} , F_{I,CO_2} , heart rate and breath-by-breath \dot{V}_{CO_2} (Fig A.1). In each case the seven curves were aligned according to the first hypoxic breath, and mean values for the 20 preceding and 70 following breaths were found on a breath-by-breath basis.

The \dot{V}_E and S_{a,O_2} curves were similar to those obtained in earlier experiments (see, for example, Fig 4.2). Between breaths 10 and 48 P_{ET,CO_2} did not differ significantly from the 20 breaths preceding the stimulus, showing that the control of CO_2 was, from the experimenter's point of view, good. The initial dip in P_{ET,CO_2} was presumably due to a more rapid increase in the subject's \dot{V}_E than expected by the experimenter so that the level of inspired CO_2 was too low for a few breaths. F_{I,CO_2} remained approximately level at 3.3% between breaths 12 and 48. The heart rate curve gave an indication of changes in cardiac output. Heart rate increased slowly from about 113 bpm to about 125 bpm over approximately 20 breaths. After an initial fluctuation, \dot{V}_{CO_2} remained at a constant level (of about 1.1 l/min compared with 1.25 l/min prior to the hypoxia) between breaths 8 and 48.

A.2.3 Discussion

During the hypoxic step, \dot{V}_{CO_2} fell to below its normoxic level and remained there until the end of the step. It is reasonable to assume that the output of CO_2 by the muscles did not change suddenly at the start of each period of hypoxia as the level of exercise remained constant throughout. This means that during the steps less CO_2 was leaving the body than was being created. If the subject's heart stroke volume remained constant, then cardiac output was proportional to heart rate, and this only rose gradually over 20 breaths. If the fall in \dot{V}_{CO_2} was due to the change in cardiac output, it would be expected that the fall would be as gradual as the rise in heart rate, but this was not the case. Instead, the fall in \dot{V}_{CO_2} reached its low level plateau at approximately the same time as inspired CO_2 reached its high level plateau. Taken together, these observations suggest that the fall in \dot{V}_{CO_2} resulted from the increase in inspired CO_2 to a level above that which was appropriate for keeping P_{a,CO_2} at a constant level, which implies that P_{a,CO_2} was increased during the period of hypoxia and so would have contributed to the stimulus to ventilation.

If it is assumed that C_{v,CO_2} did not decrease significantly with the increase in cardiac output, which is quite possible given the modest increase in cardiac output and the high CO_2 capacity of the tissues, then Equation 1 indicates that \dot{V}_{CO_2} should be **higher** for a higher level of cardiac output if P_{a,CO_2} is to remain constant. The fact that it was lower in practice adds further weight to the suggestion that P_{a,CO_2} was being raised by the addition of inspired CO_2 .

The conclusion from this experiment would appear to be that the action taken to prevent P_{a,CO_2} falling during hypoxia can result in P_{a,CO_2} being raised above its normoxic level. Rather than removing a potential factor affecting the ventilatory response, this method might instead be adding to the stimulus both at the carotid body and centrally.

Unfortunately there was insufficient time to pursue this line of experiment further. The \dot{V}_{CO_2} algorithm in the data acquisition program needs to be checked independently by comparison with \dot{V}_{CO_2} measured using a standard gas collection method; and results for several more subjects should be obtained. As this was not possible, the matter was investigated further using a simple model.

A.3 A MODEL OF THE P_{A,CO_2} CYCLE DURING BREATHING

A.3.1 Introduction

Alveolar P_{CO_2} (P_{A,CO_2}) varies cyclically as lung volume changes during breathing. Its waveform was first calculated by DuBois et al (1952) without the benefit of a computer, using an elegant but laborious method. In order to investigate the changing difference between P_{ET,CO_2} and mean P_{a,CO_2} (assumed equal to mean P_{A,CO_2}), I implemented their model on a microcomputer and extended it to include the ability to simulate the method of keeping P_{ET,CO_2} constant by adding inspired CO_2 , simulation of the dead space associated with the respiratory valve, calculations for \dot{V}_{CO_2} , and an interface to Lotus-123 files that allowed complex series of experiments to be set up.

A.3.2 The Model

The model of DuBois et al makes several assumptions: mixing at the alveolar level is instantaneous, but no diffusion occurs in the deadspace gas; P_{a,CO_2} and P_{A,CO_2} are always assumed equal (Suskind et al, 1950), and DuBois et al use a concept of "equivalent lung volume for CO_2 " which replaces lung volume throughout. The equivalent lung volume is the value found during breath-holding experiments designed to estimate lung volume. These experiments resulted in values for lung volume of about 5 litres, presumably due to the effects of CO_2 being absorbed into and released from lung tissues and blood capillaries, which would cause the lung alveoli to appear to have an increased store of CO_2 .

DuBois et al started with assumed values for lung volume, dead space volume, P_{v,CO_2} , cardiac output and initial P_{A,CO_2} , and then calculated values for P_{A,CO_2} every 0.1 seconds during a whole breath cycle using a breath pattern obtained from experimental data. Each P_{A,CO_2} value was obtained from the previous by calculating the new total quantity of CO_2 in the alveoli and then dividing by the new volume. The quantity of CO_2 changed through three mechanisms:

1. CO_2 was delivered by the blood according to the formula:

$$\Delta P_{A,CO_2} = \frac{(P_{v,CO_2} - P_{A,CO_2})(\text{dissoc.slope})(\dot{Q})(B-47)}{\text{eq.lung.vol}} \Delta t$$

(where "dissoc. slope" is the slope of the appropriate part of the CO_2 dissociation curve, B-47 is atmospheric pressure less water vapour pressure, and "eq.lung.vol" is equivalent lung volume).

2. CO_2 left the lung with the outgoing air during expiration.
3. During early inspiration CO_2 was also rebreathed from the deadspace: this quantity of CO_2 was calculated from stored values of P_{A,CO_2} from the end of the previous expiration.

To this simple model I added a representation of the respiratory valve and mass spectrometer probe. The former increased the total dead space volume, and therefore the amount of CO_2 to be rebreathed during inspiration. The latter measured the P_{CO_2} at the "mouth", ie between the

tracheal deadspace and the valve. Dead space was simulated by storing in an array pairs of values for volume and P_{CO_2} expired during each sample interval. The P_{CO_2} at the mouth was obtained by summing the expired volumes from the "lung" end of the array until a total volume equivalent to the volume of the dead space was reached, and then reading off the associated value of P_{CO_2} (using linear interpolation to increase precision). P_{ET,CO_2} was taken as the value of P_{CO_2} at the mouth at the end of expiration. \dot{V}_{CO_2} was calculated from the P_{CO_2} of gas entering and leaving the mouth and the volume inspired or expired, dividing the total volume of CO_2 leaving the mouth by the period of the breath.

The model also allowed the simulation of inspired CO_2 . After the dead space gas had been inspired, the inspired P_{CO_2} was set to the level required, and the calculations proceeded exactly as before. A simple proportional feedback controller was added to allow inspired CO_2 to be controlled in such a way as to keep P_{ET,CO_2} at a pre-defined level. The controller set the level of inspired CO_2 once per breath.

Rather than use an experimentally-obtained breathing pattern as was used by DuBois et al, the breathing pattern was simulated using segments of sine waves with a chosen amplitude (V_T) added to a basic volume (frc) according to the formula:

$$\begin{aligned} \text{Inspiration:} \quad V &= \frac{V_T}{2} \left(1 - \cos \frac{t}{T_I} \pi \right) + \text{frc} \\ \text{Expiration:} \quad V &= \frac{V_T}{2} \left(1 + \cos \frac{t-T_I}{T_E} \pi \right) + \text{frc} \\ \text{Pause:} \quad V &= \text{frc} \end{aligned}$$

This method allowed inspiratory, expiratory and end-expiratory pause times to be set independently.

The model was implemented in GW-BASIC on an IBM compatible PC. Control of variables such as V_T , T_I and cardiac output was accomplished by reading them from a Lotus-123-produced file. This allowed complex series of experiments to be set up and run. Values of P_{ET,CO_2} , mean P_{A,CO_2} , \dot{V}_{CO_2} etc. were stored on a breath-by-breath basis, while P_{CO_2} at the mouth

and P_{A,CO_2} were stored at 10 values per simulated second, the sampling interval for calculations.

A.3.3 Experiment 1. Factors affecting the P_{ET,CO_2} - P_{A,CO_2} difference

Values for V_T , f_B , P_{v,CO_2} and cardiac output corresponding to a \dot{V}_{O_2} of 1 l/min were found (using interpolation) from the guide to normal values in Jones et al (1975). V_T , T_I , T_E , T_{pause} , inspired CO_2 , P_{v,CO_2} and cardiac output were varied in turn to see the effect of each on the difference between P_{ET,CO_2} and mean P_{A,CO_2} . Results were stored once equilibrium had been reached for each set of values.

A.3.4 Results:

Table A.1 shows the results for each variable. The difference between P_{ET,CO_2} and mean P_{A,CO_2} increased with

- tidal volume,
- breath period,
- expiratory time,
- cardiac output;

and decreased with

- breathing frequency,
- inspired CO_2 ,
- increasing ratio of $T_I:T_E$,
- the existence of an end-expiratory pause.

A.3.5 Experiment 2. Keeping P_{ET,CO_2} constant using inspired CO_2

During step hypoxia V_T and breathing frequency both increase. According to the results of Experiment 1, the rise in V_T will increase the [ET-a] difference whilst the increase in breathing frequency will decrease it. The difference may also be reduced by CO_2 inspired during the step. It is

important to establish whether the overall result is an increase or decrease in the [ET-a] difference.

This experiment started with 10 breaths using the same parameter values as at the start of Experiment 1 to obtain a "normoxic" value for P_{ET,CO_2} , and then \dot{V}_E was increased to twice its initial value with different combinations of V_T and frequency, for each of which P_{ET,CO_2} was brought back to its normoxic level by changing the level of inspired CO_2 .

A.3.6 Results:

Table A.2 shows that the [ET-a] difference depended heavily on the combination of V_T and frequency even though \dot{V}_E and P_{ET,CO_2} were constant. The difference, as in Experiment 1, increased if V_T alone increased, though not by as much as would be predicted from an extrapolation of the Experiment 1 results because of the addition of inspired CO_2 . As breathing frequency increased (ie $T_I + T_E$ decreased) the difference rapidly shrank to below its initial value. \dot{V}_{CO_2} also decreased to below its initial level, just as in the study involving breath-by-breath \dot{V}_{CO_2} calculations described earlier in this appendix.

Table A.1 Factors affecting the [ET-a] difference.

The basic set of parameter values used was

Tidal volume (V_T)	=	1.4 l
Inspiratory time (T_I)	=	1.4 s
Expiratory time (T_E)	=	1.8 s
End-exp. pause time	=	0 s
Inspired P_{a,CO_2}	=	0 kPa
Venous P_{a,CO_2} (P_{v,CO_2})	=	8.2 kPa
Blood flow (\dot{Q})	=	0.181/s

cont .../

These were varied in turn, and model values were found for P_{ET,CO_2} and mean P_{a,CO_2} (allowing 20 breaths after each change to reach steady-state).

(Note: table continued on next page)

	P_{ET,CO_2} (kPa)	P_{a,CO_2} (kPa)	[ET-a] (kPa)
Parameter values as above	5.29	5.24	.05
$V_T=1.0$	6.03	6.03	.00
1.4	5.29	5.24	.05
1.8	4.84	4.76	.08
2.2	4.28	4.16	.12
$T_I=1.0$	5.14	5.08	.06
1.4	5.29	5.24	.05
2.0	5.61	5.58	.03
$T_E=1.8$	5.29	5.24	0.05
2.2	5.53	5.46	0.07
Pause =0.0	5.29	5.24	.05
0.2	5.57	5.56	.00
0.5	5.29	5.26	.03
$T_I=1.0$ $T_E=1.3$	4.74	4.72	.02
1.4 1.8	5.29	5.24	.05
2.0 2.6	5.97	5.90	.07
Insp $CO_2=0.0$	5.29	5.24	.05
1.0	5.65	5.61	.04
2.0	6.00	5.96	.04
3.0	6.36	6.33	.03

$V_T = 1.8$ Insp $CO_2 = 1.0$	5.26	5.19	0.07
2.0	5.66	5.60	0.06
3.0	6.07	6.02	0.05
$P_{v,CO_2} = 7.7$	4.98	4.94	.04
8.2	5.29	5.24	.05
8.5	5.48	5.43	.05
$Q = 0.15$	4.94	4.90	0.04
0.18	5.29	5.24	0.05
0.21	5.58	5.52	0.06

Table A.2 Variation of the [ET-a] difference despite constant \dot{V}_E . Row **a** shows model P_{ET,CO_2} , mean P_{a,CO_2} and the difference [ET-a] for the basic set of values given in Table A.1. Rows b-h are all for a value of \dot{V}_E twice that in Row a, but with different V_T and breath periods. Inspired P_{CO_2} was adjusted to bring P_{ET,CO_2} to the level of that in Row a. "Factor" is the [ET-a] difference scaled so that Row a has the value 1.0.

Row	V_T	T_I	T_E	Pause	P_{CO_2} ET	P_{CO_2} a	[ET-a]	factor
	(l)	(s)	(s)	(s)	(kPa)	(kPa)	(kPa)	
a	1.40	1.400	1.800	0.0	5.295	5.247	0.048	1.000
b	2.80	1.400	1.800	0.0	5.295	5.196	0.099	2.082
c	2.45	1.225	1.575	0.0	5.295	5.226	0.069	1.453
d	2.10	1.050	1.350	0.0	5.295	5.243	0.053	1.106
e	1.75	0.875	1.125	0.0	5.295	5.263	0.032	0.677
f	1.40	0.700	0.900	0.0	5.294	5.290	0.004	0.082
g	2.10	1.350	1.050	0.0	5.295	5.271	0.024	0.507
h	2.10	1.050	1.050	0.3	5.295	5.301	-.006	-.125

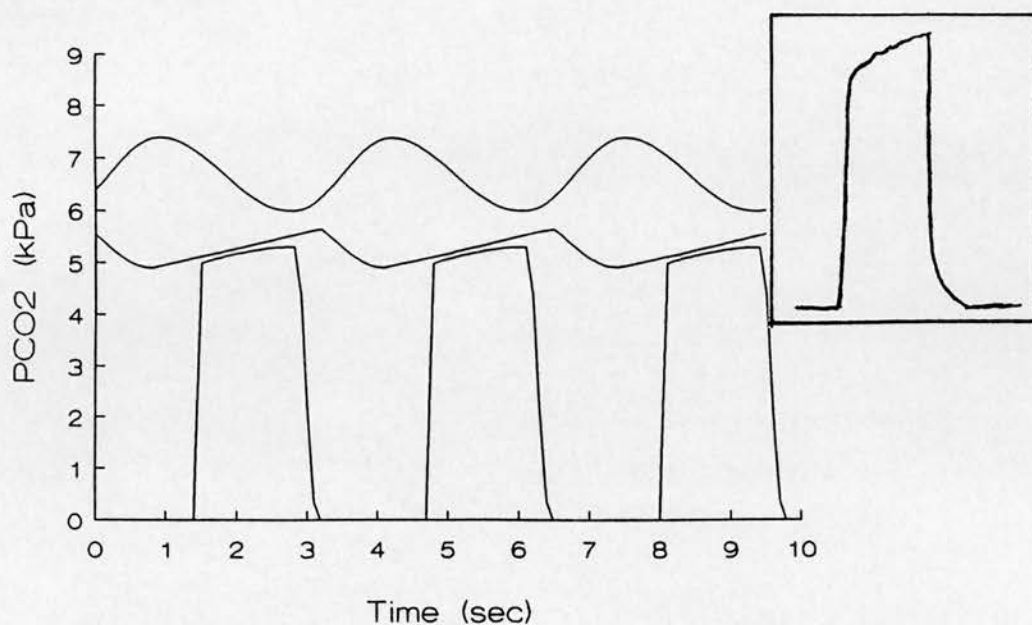
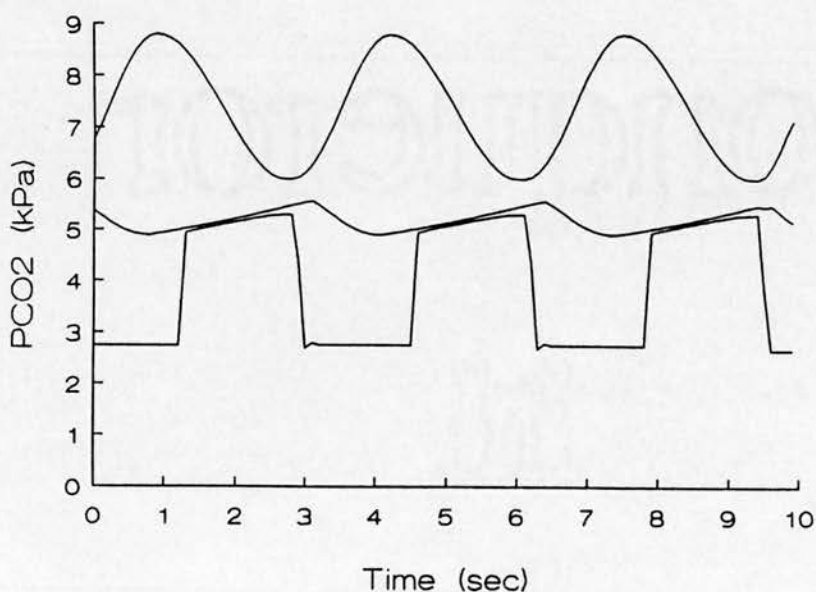


Fig A.2 Model output. *a)* the output corresponding with Row *a* in Table A.2, with (inset) an experimentally obtained P_{CO_2} curve for a similar level of exercise for comparison. The lower curve is P_{CO_2} measured at the mouth, the middle curve is the alveolar P_{CO_2} curve, and the upper curve is the lung volume curve for reference (plotted in arbitrary units).



b) Output when V_T increased to double that in *a*).

cont.../

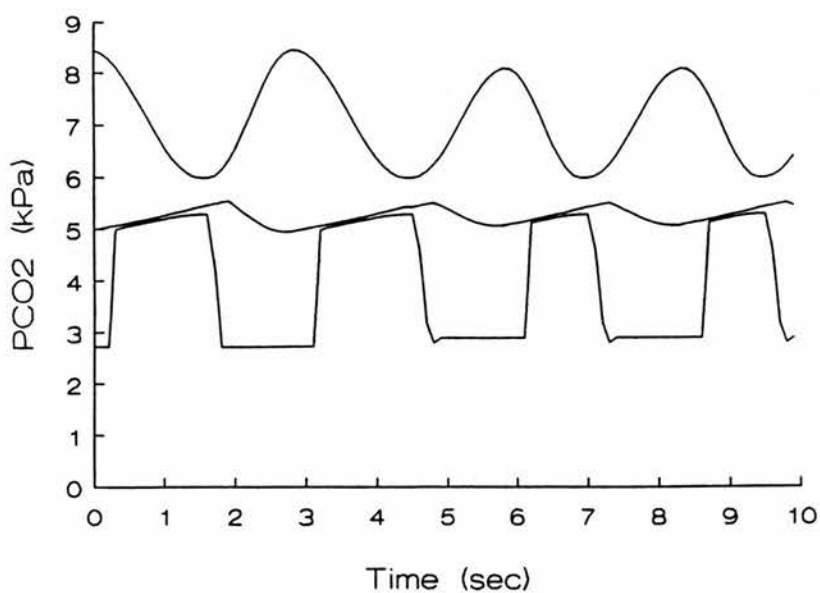
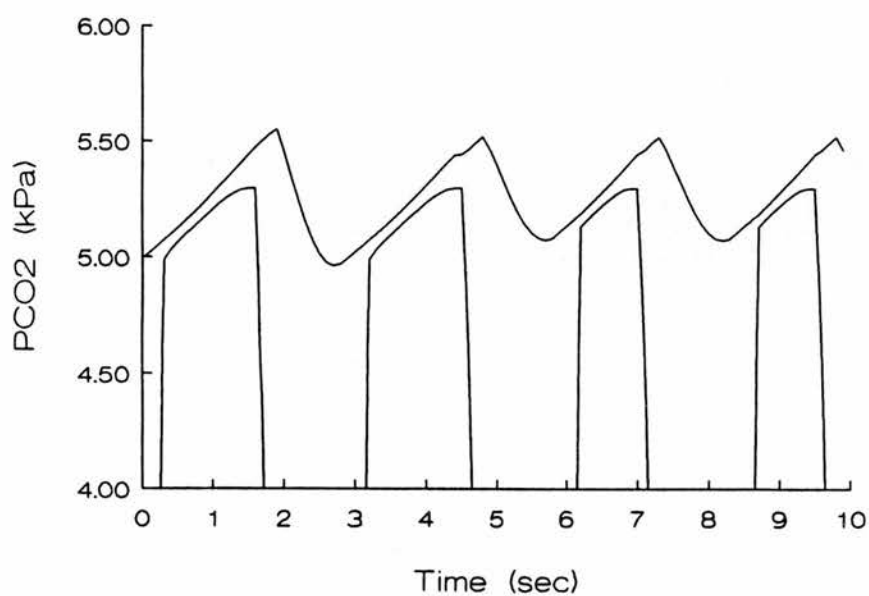


Fig A.2 c) Outputs for two different frequencies but with the same \dot{V}_E as in b).



d) Repeat of c) but with increased vertical axis scale for clarity.

A.4 DISCUSSION

Fig A.2 shows sample oscillations in P_{A,CO_2} and the resulting trace of P_{CO_2} measured as if at the mouth. The P_{A,CO_2} oscillations were similar to those calculated by DuBois et al. Their shape was roughly saw-toothed, with P_{A,CO_2} rising from shortly before the start of expiration (when the flow of fresh air entering the lungs was less than the flow of CO_2 from the blood) to a time well after the start of inspiration (due to the rebreathing of deadspace gas). The curvature of the oscillations depended on the shape of the flow profile of gas entering and leaving the lung. The expired P_{CO_2} pattern was qualitatively similar to that observed, except for certain details. Diffusion in the trachea led to a less sharp rise in P_{CO_2} at the start of expiration than seen in the model output. The fall in P_{CO_2} at the start of inspiration is generally a little steeper than calculated by the model: this is probably because the model's flow profile, based on sine wave segments, was unrealistic at this point, with inspired flow starting less rapidly than observed in exercise studies. More significantly, the slope of the top section of the curve was less steep than observed on exercise (see inset to Fig A.2a). This could be due to a number of reasons. The "equivalent lung volume" of the DuBois et al model may require alteration to be valid on exercise. Certainly, a reduction in value from 5 litres to 1 litre produced a much more realistic slope. A second possibility is that the difference is due to gas mixing in the lungs not being instantaneous. Expired gas from early expiration would then contain a disproportionately large amount of the inspired gas, and so would have a lower P_{CO_2} than the average of the lung as a whole. A further reason might be that the simple equation dictating the quantity of CO_2 delivered by the blood may be incomplete in some way.

The experiments outlined above were repeated with an "equivalent lung volume" of 1 litre, in case this was the cause behind the discrepancy in P_{CO_2} slope. The values of [ET-a] were all very much increased (by a factor of 2 to 4) but all the trends in the results remained the same. The [ET-a] difference was very sensitive to the flow profile used, and so using the artificial flow profile based on sine waves rather than a more realistic shape meant that the absolute values produced were unlikely to be

meaningful in themselves. The trends in the results were therefore the parts of most value.

If the model is valid, then the difference between P_{ET,CO_2} and mean P_{a,CO_2} varies with breathing pattern. The difference increases with tidal volume, expiratory time and cardiac output, and decreases with breathing frequency, the level of inspired CO_2 and the existence of an end-expiratory pause. The difference is affected by the shape of the flow profile and the division of the breath period between inspiration and expiration. When P_{ET,CO_2} is kept constant by adding CO_2 to the inspired gas, the difference still varies with tidal volume and frequency even though \dot{V}_E might be constant (Fig A.2 shows example oscillations). This implies that keeping P_{ET,CO_2} constant using inspired CO_2 is likely to cause a systematic change in mean P_{a,CO_2} . Unfortunately, so many factors seem to affect the difference that it would be very hard to predict what the difference should be at any particular time.

What can be done to minimise the problem? Jones et al (1979) proposed a formula obtained from their data by multiple regression analysis. This relates P_{a,CO_2} to P_{ET,CO_2} and V_T and so can be used to obtain values for P_{a,CO_2} . There are several problems with using this in hypoxia studies. Firstly, the formula does not take into account many of the possible factors affecting the [ET-a] difference. (It is interesting to note that their data seems to support the observation that the difference is dependent on frequency as well as V_T , and so it may be possible to derive a more detailed formula. Frequency features in a similar regression equation derived in an earlier study along the same lines (Jones et al, 1966).) Secondly, the formula was derived using the data from only 5 healthy male subjects, and so may not be applicable to a general population. Indeed, for one of those 5 subjects (Subject 1), P_{ET,CO_2} provided a better indicator of P_{a,CO_2} than the value obtained from the formula. Thirdly, the formula is very likely to be inappropriate when CO_2 is added to the inspired gas mixture. Describing such circumstances was not one of the aims of the authors. It may be possible to repeat the work of Jones et al, including values for P_{a,CO_2} measured during hypoxic steps and considering a wider variety of possible factors in the regression analysis. Inter-subject variability may still limit the usefulness of such formulae.

An alternative approach might be to perfect the model and adapt it so that it could be used online during a study, sampling the flows and mass spectrometer output and using these to estimate P_{a,CO_2} . Control of the CO_2 flow might even be automated as part of the program. It may be possible, if the model were extended, to build in a method of predicting the required level of inspired CO_2 part way through a breath, allowing better control of P_{a,CO_2} .

A further alternative might be to develop a method based round the idea of keeping breath-by-breath \dot{V}_{CO_2} at an appropriate level. This would involve keeping track of changes in cardiac output (ie in heart rate as long as stroke volume remains constant) and calculating the change in \dot{V}_{CO_2} that should accompany this. CO_2 would then be added to the inspired gas to keep \dot{V}_{CO_2} at the correct level.

The work of this chapter is just a start. Much more needs to be done. The model is crude, and could be improved by considering slower gas mixing in the alveoli (which could potentially have a big impact on the results), a better idea of how "equivalent lung volume" relates to actual lung volume (to allow the latter, which can be measured using standard techniques, to be used directly), and using more realistic flow profiles. The experimental work described earlier in this appendix needs to be expanded and checked. Yet despite these limitations, the two parts of this appendix suggest that there is a strong possibility that the standard method for controlling CO_2 during hypoxia is fundamentally flawed. Any experiment which involves the assumption that P_{ET,CO_2} is equal to mean P_{a,CO_2} is likely to introduce a systematic change in the level of P_{a,CO_2} , and therefore in the stimulus to the carotid body and central receptor. If this is the case, then a lot of work in this field, including that from our own laboratory, must be viewed with considerable care.

REFERENCES

- ⁽¹⁹⁷⁹⁾
ANTONIOU A. Digital Filters: Analysis and Design. McGraw-Hill.
- BAND DM, McCLELLAND M, PHILLIPS D, SAUNDERS K, WOLFF CB.** Sensitivity of the carotid body to within breath changes in arterial PCO₂. *J Appl Physiol* 1978; 45:768-777.
- BARTLETT D Jr, TENNEY SM.** Control of breathing in experimental anaemia. *Resp Physiol* 1970; 10:384-395.
- BELLVILLE JW, WHIPP BJ, KAUFMAN RD, SWANSON GD, AQLEH KA, WIBERG DM.** Central and peripheral chemoreflex loop gain in normal and carotid body-resected subjects. *J Appl Physiol* 1979; 46:843-853.
- BISCOE TJ, PURVES MJ.** Observations on the rhythmic variation in the cat carotid body chemoreceptor activity which has the same period as respiration. *J Physiol* 1967; 190: 389-412.
- BRADLEY CA, FLEETHAM JA, ANTHONISEN NR.** Ventilatory control in patients with hypoxaemia due to obstructive lung disease. *Am Rev Resp Dis* 1979; 120:21-29.
- CARSON ER, COBELLI C, FINKELSTEIN L.** The Mathematical Modeling of Metabolic and Endocrine Systems, p64. Wiley and Sons, New York (1983).
- CARSON ER, MURRAY-SMITH DJ** The conceptual basis of modelling the respiratory system. In: *The Respiratory System* (Cramp DG, Carson ER eds). pp296-333. Croom Helm, London and Sydney (1988).
- CHERNIAK NS, EDELMAN NH, LAHIRI S.** Hypoxia and hypercapnia as respiratory stimulants and depressants. *Resp Physiol* 1971; 11:113-126.
- COLLINS DD, SCOGGIN CH, ZWILLICH CW, WEIL JV.** Hereditary aspects of decreased hypoxic response. *J Clin Invest* 1978; 62:105-110.
- COMROE JH Jr.** The location and function of the chemoreceptors of the aorta. *Am J Physiol* 1939; 127:176-191.

COMROE JH Jr, SCHMITT CF. The part played by reflexes from the carotid body in the chemical regulation of respiration in the dog. *Am J Physiol* 1938; 121:75-97.

CRAMP DG, CARSON ER. Eds. *The Respiratory System*. Croom Helm, London and Sydney (1988).

CROSS BA, LEAVER KD, SEMPLÉ SJG, STIDWILL RP. Response of carotid body chemoreceptor discharge to small changes in PaCO₂ in the cat. *J Physiol* 1984; 353:131P.

CROSS KS, OPPE TW. The effect of high and low concentrations of oxygen on the respiration of the premature infant. *J Physiol (Lond)* 1952; 117:38-55.

CROSS BA, SEMPLÉ SJG. Do oscillations in arterial CO₂ tension provide feed-forward control of ventilation? In: *The Control of Breathing in Man*, Whipp BJ Ed; pp68-86. Manchester University Press, Manchester (1987).

DALY MDeB, LAMBERTSON CJ, SCHWEITZER A. Observations on the volume of blood flow and oxygen utilisation of the carotid body in the cat. *J Physiol (Lond)* 1954; 125:67-89.

DAUBENSPECK JA. Frequency analysis of CO₂ regulation: Afferent influence on tidal volume control. *J Appl Physiol* 1973; 35:662-672.

De CORT SC. Measurement of hypoxic ventilatory drive at rest and during exercise in normal man. PhD Thesis, University of Edinburgh, 1989.

DEFARES JC, DERKSEN HE, DUYFF JW. Cerebral blood flow in the regulation of respiration. *Acta Physiol Pharmacol Neerl* 1960; 9:327-360.

DEJOURS P. Control of respiration in muscular exercise. In: *Handbook of Physiology* (Fenn WO & Rahn H eds); pp 631-48. American Physiological Society 1964, Washington DC.

DEJOURS P. La regulation de la ventilation au cours de l'exercice musculaire chez l'homme. *J Physiol (Paris)* 1953; 51:163-261.

DOUGLAS NJ, BRASH HM, WRAITH PK, CALVERLEY PMA, LEGGETT RJE, McELDERRY L, FLENLEY DC. Accuracy, sensitivity to

carboxyhaemoglobin and speed of response of the Hewlett-Packard 47201A ear oximeter. *Am Rev Res Dis* 1979; 119:311-313.

DRIPPS RD, COMROE JH Jr. The effect of the inhalation of high and low oxygen concentrations on respiration, pulse rate, ballistocardiogram and arterial oxygen saturation (oximeter) of normal individuals. *Am J Physiol* 1947; 149:277.

DuBOIS AB, BRITT AG, FENN WO. Alveolar CO₂ during the respiratory cycle. *J Appl Physiol* 1952; 4:535-548.

DUFFIN J. A mathematical model of the chemoreflex control of ventilation. *Respir Physiol* 1972; 15:277-301.

DUKE HN, GREEN JH, NEIL E. Carotid chemoreceptor impulse activity during inhalation of carbon monoxide mixture. *J Physiol (Lond)* 1952; 118:520-527.

DUTTON RE, SMITH EJ, GHATAK PK, DAVIES DG. Dynamics of the respiratory controller during carotid body hypoxia. *J Appl Physiol* 1973; 35:844-850.

EASTON PA, SLYKERMANN LJ, ANTHONISEN NR. Recovery of the ventilatory response to hypoxia in normal adults. *J Appl Physiol*. 1988; 64(2):521-528.

EASTON PA, SLYKERMANN LJ, ANTHONISEN NR. Ventilatory response to sustained hypoxia in normal adults. *J Appl Physiol* 1986; 61(3):906-911.

EDELMAN NH, EPSTEIN PE, LAHIRI S, CHERNIACK NS. Ventilatory response to transient hypoxia and hypercapnia in man. *Respir Physiol* 1973; 17:302-314.

ELDRIDGE FL, MILLHORN DE, KILEY JP, WALDROP TG. Stimulation by central command of locomotion, respiration and circulation during exercise. *Respir Physiol* 1985; 59:313-337.

EYZAGUIRRE C, LEWIN J. Chemoreceptor activity of the carotid body of the cat. *J Physiol (Lond)* 1961; 159:222-237.

FIDONE SJ Putative neurotransmitters on the carotid body: the case of dopamine. *Fed Proc* 1980; 39:2636-2640.

FLENLEY DC, BRASH H, CLANCY L, COOKE NJ, LEITCH AG, MIDDLETON W, WRAITH PK. Ventilatory response to steady-state exercise in hypoxia in humans. *J Appl Physiol* 1979; 46(3):438-446.

FLENLEY DC, FRANKLIN DH, MILLER JS. The hypoxic drive to breathing in chronic bronchitis and emphysema. *Clin Sci* 1970; 38:503-518.

FLENLEY DC, MILLER JS. Ventilatory response to transient hypoxia and hypercapnia in man. *Resp Physiol* 1967; 17:302-314.

FORSTER HV, DEMPSEY JA, BIRNBAUM ML, REDDAN WG, THODEN J, GROVER RF, RANKIN J. Effect of chronic exposure to hypoxia on ventilatory response to CO₂ and hypoxia. *J Appl Physiol* 1971; 31(4):586-592.

GANDEVIA SC, MACEFIELD G. Projection of low-threshold afferents from human intercostal muscles to the cerebral cortex. *Resp Physiol* 1989; 77:203-214.

GARDNER WM. The pattern of breathing following step changes of alveolar partial pressures of carbon dioxide and oxygen in man. *J Physiol* 1980; 300:55-73

GOODMAN NW, NAIL BS, TORRANCE RW. Oscillations in the discharge of single carotid chemoreceptor fibres in the cat. *Resp Physiol* 1974; 20:251-263.

GOULD GA, AIRLIE MAA, BRASH HM, WRAITH PK, WARREN PM, FLENLEY DC. The "On-phase" of ventilatory response to transient hypoxia to assess carotid chemoreceptor mediated ventilatory drive during exercise. *Clin Sci* 1985; 68 suppl 11:62P.

GRAY JS. The multiple factory theory of the control of respiratory ventilation. *Science* 1946; 102:739-744.

GRIFFITHS TL, HENSON LC, WHIPP BJ. The influence of inspired O₂ concentration on the dynamics of the exercise hyperpnoea in man. *J Physiol* 1986; 380:387-403.

GRODINS FS, BUELL J, BART AJ. Mathematical analysis and digital simulation of the respiratory control system. *J Appl Physiol* 1967; 22:260-276.

GRODINS FS, GRAY JS, SCHROEDER KR, NORINS AL, JONES RW. Respiratory responses to CO₂ inhalation: A theoretical study of a nonlinear biological regulator. *J Appl Physiol* 1954; 7:283-308.

GRODINS FS, JAMES G. Mathematical models of respiratory regulation. *Ann N Y Acad Sci* 1963; 109:852-868.

GRODINS FS, YAMASHIRO SM. Control of ventilation. In: *Regulation of Breathing Part II* (Hornbein TF ed) Marcel Dekker New York 1981; pp515-558.

GUZ A, NOBLE MIM, WIDDICOMBE JG, TRENCHARD D, MUSHIN WW. Peripheral chemoreceptor block in man. *Resp Physiol* 1966; 1:38-40.

HALDANE JS, PRIESTLEY JG. The regulation of the lung ventilation. *J Physiol (Lond)* 1905; 32:225-266.

HEBBEL RP, KRONENBERG RS, EATON JW. Hypoxic ventilatory response in subjects with normal and high oxygen affinity haemoglobins. *J Clin Invest* 1977; 60:1211-1215.

HILL JE. The dynamic hypoxic ventilatory response in normal subjects and its contribution to the severity of hypoxaemia in chronic obstructive pulmonary disease. PhD Thesis, University of Edinburgh, 1990.

HIRSHMAN CA, MCCULLOUGH RE, WEIL JV. Normal values for hypoxic and hypercapnic ventilatory drives in man. *J Appl Physiol* 1975; 38:1095-1098.

HOLTON P, WOOD JB. The effects of bilateral removal of the carotid bodies and denervation of the carotid sinuses in two human subjects. *J Physiol (Lond)* 1965; 181:365-378.

HONDA Y, WATENABE S, HASHIGUME I, SATOMURA Y, HATA N, SAKAKIBARA Y, SEVERINGHAUS JW. Hypoxic chemo-sensitivity in asthmatic patients two decades after carotid body resection. *J Appl Physiol* 1979; 46:632-638.

HORGAN JD, LANGE RL. Digital computer simulation of the human respiratory system. *IEEE International Conference Record* 1963. pp149-157.

HORGAN JD, LANGE RL. Digital computer simulation of respiratory response to cerebrospinal fluid PCO₂ in the cat. *Biophys J* 1965; 6:935-945.

HORNBEIN TF. The relation between the stimulus to chemoreceptors and their response. In: *Arterial Chemoreceptors*. Ed R W Torrance, Oxford, England. Blackwell Scientific Publications 1968; pp 65-78.

HORNBEIN TF, GRIFFO ZJ, ROOS A. Quantitation of chemoreceptor activity: interrelation of hypoxia and hypercapnia. *J Neurophysiol* 1961; 24:561-568.

JAIN SK, SUBRAMANIAN S, JULKA DB, GUZ A. Search for evidence of lung chemoreflexes in man: study of respiratory and circulatory effects of phenyldiguanine and lobeline. *Clin Sci* 1972; 42:163-177.

JONES NL, CAMPBELL EJM, EDWARDS RHT, ROBERTSON DG. *Clinical Exercise Testing*. p54,p205. WB Saunders Co, Philadelphia, London, Toronto (1975).

JONES NL, MCHARDY GJR, NAIMARK A, CAMPBELL EJM. Physiological dead space and alveolar-arterial gas pressure differences during exercise. *Clin Sci* 1966; 31:19-29.

JONES NL, ROBERTSON DG, KANE JW. Difference between end-tidal and arterial Pco₂ in exercise. *J Appl Physiol*. 1979; 47(5):954-960.

KAO FF. An experimental study of the pathways involved in exercise hyperapnea employing cross circulation techniques. In: *The Regulation of Human Respiration*, Cunningham DJC & Lloyd BB Eds; pp461-502. Blackwell, Oxford (1963).

- KAWAKAMI Y, IRIE T, SHIDA A, YOSHIKAWA T.** Familial factors affecting arterial blood gas values and respiratory chemosensitivity in chronic obstructive pulmonary disease. *Am Rev Resp Dis* 1982; 125:420-425.
- KHOO MCK, YAMASHIRO SM.** Models of control of breathing. In: *Respiratory Physiology: An Analytical Pproach*; Chang HK and Paiva M Eds; Dekker, New York pp799-829 (1989).
- KROGH A, LINDHARD J.** The regulation of respiration and circulation during the initial stages of muscular work. *J Physiol* 1913; 47:112-136.
- KRONENBERG RS, DRAGE CW.** Attenuation of the ventilatory and heart rate responses to hypoxia and hypercapnia with aging in normal man. *J Clin Invest* 1973; 52:1812-1819.
- KRONENBERG R, HAMILTON FN, GABEL R, HICKEY R, READ DJC, SEVERINGHAUS JW.** Comparison of three methods for quantitating respiratory response to hypoxia in man. *Resp Physiol* 1973; 16:109-125.
- LAHIRI S, GELFAND R.** Mechanisms of acute ventilatory responses. In: *Regulation of Breathing, part II* (Hornbein TF ed), Dekker, New York, 1981; pp773-843.
- LAKSHMINARAYAN S, SAHN SA, WEIL JV.** Effect of aminophylline on ventilatory response in normal man. *Am Rev Resp Dis* 1978; 117:33-38.
- LAUBIE M, SCHMITT H.** Long-lasting hyperventilation induced by almitrine: evidence for a specific effect on carotid and thoracic chemoreceptors. *Eur J Pharmacol* 1980; 61:125-136.
- LEITCH AG, CLANCY L, FLENLEY DC.** Maximal oxygen uptake, lung volume and ventilatory response to carbon dioxide and hypoxia in a pair of identical twin athletes. *Clin Sci Mol Med* 1975; 48:235-238.
- LLOYD BB, CUNNINGHAM DJC.** Quantitative approach to the regulation of human respiration. In: *The Regulation of Human Respiration*, (Cunningham DJC & Lloyd BB eds.) Blackwell, Oxford 1963; pp 331-349.

LLOYD BB, JUKES MGM, CUNNINGHAM DJC. The relation between alveolar oxygen pressure and the respiratory response to carbon dioxide in man. *Q J Exp Physiol* 1958; 43:214-227.

LONGOBARDO GS, CHERNIACK NS, FISHMAN AP. Cheyne-Stokes breathing produced by a model of the human respiratory system. *J Appl Physiol* 1966; 21:1839-1846.

LONGOBARDO GS, GOTHE B, GOLDMAN MD, CHERNIACK NS. Sleep apnea considered as a control system instability. *Resp Physiol* 1982; 50:311-333.

LUGLIANI R, WHIPP BJ, SEARD C, WASSERMAN K. Effect of bilateral carotid resection on ventilatory control at rest and during exercise in man. *N Eng J Med* 1971; 285:1105-1111.

MARGENAU H, MURPHY GM. The Mathematics of Physics and Chemistry, 2nd Ed, pp 467-519. D Van Nostrand Co, London (1956).

MATSUURA S. Chemoreceptor properties of glomus tissue found in the carotid region of the cat. *J Physiol (Lond)* 1973; 235:57-63.

McCLOSKEY EI, MITCHELL JH. Reflex cardiovascular and respiratory responses originating in exercising muscle. *J Physiol* 1972; 224:173-186.

McDONALD DM. Peripheral chemoreceptors: structure function relationships in the carotid body. In *Regulation of Breathing, Part I* (Hornbein TF ed), Dekker, New York 1981; pp 105-319.

MILLHORN HT Jr, REYNOLDS WJ. Digital simulation of the chemical control of ventilation. *Trans ASME: J Dynamic Systems* 1973; Measure Control 95(Ser G):335-339.

MILLHORN HT Jr, BENTON R, ROSE R, GUYTON AC. A mathematical model of the human respiratory control system. *Biophys J* 1965; 5:27-46.

MILLHORN HT Jr, REYNOLDS WJ, HOLLOMAN GH Jr. Digital simulation of the ventilatory response to CO₂ inhalation CSF perfusion. *Comput Biomed Res* 1972; 5:301-314.

MURRAY-SMITH DJ, CARSON ER. The modelling process in respiratory medicine. In: *The Respiratory System* (Cramp DG, Carson ER eds). pp296-333. Croom Helm, London and Sydney (1988).

NELDER JA, MEAD R. *Computer Journal* 1965; 7:308.

NIELSEN M, SMITH H. Studies on the regulation of respiration in acute hypoxia. *Acta Physiol Scand* 1951; 24:293-313.

NISHIMURA M, SUZUKI A, NISHIURA Y, YAMAMOTO H, MIYAMOTO K, KISHI F, KAWAKAMI Y. Effect of brain blood flow on hypoxic ventilatory response in humans. *J Appl Physiol* 1987; 63: 1100-1106.

PATIL CP, CUMMIN ARC, JACOBI MS, IYAWA VI, SAUNDERS KB. Ensemble-averaging of breath-to-breath data for CO₂ responses. *Clin Sci* 1984; 67:60P.

PETERSEN ES, VEJBY-CHRISTENSEN H. Effects of body temperature on ventilatory response to hypoxia and breathing pattern in man. *J Appl Physiol* 1977; 42:492-500.

PETERSON DD, PACK AI, SILAGE DA, FISHMAN AP. Effects of aging on ventilatory and occlusion pressure responses to hypoxia and hypercapnia. *Am Rev Resp Dis* 1981; 124:387-450.

POON CS. Optimal control of ventilation in hypoxia, hypercapnia and exercise. In: *Modeling and Control of Breathing* (Whipp BJ, Wiberg DM eds) pp189-196, Elsevier, New York (1983).

PRESS WH, FLANNERY BP, TEUKOLSKY SA, VETTERLING WT. *Numerical Recipes. The Art of Scientific Computing*, pp274-334, 390-429, 459, 547-577. Cambridge University Press, Cambridge (1986).

REBUCK AS, CAMPBELL EJM. A clinical method for assessing the ventilatory response to hypoxia. *Am Rev Resp Dis* 1974; 109:345-350.

REYNOLDS WJ, MILHORN HT. Transient ventilatory response to hypoxia with and without controlled alveolar Pco₂. *J Appl Physiol* 1973; 35:187-196.

RIGATTO H, BRADY JP, de la TORRE VERDUZCO R. Chemoreceptor reflexes in preterm infants: I The effect of gestational and postnatal age on the ventilatory response to inhalation of 100% and 15% oxygen. *Pediatrics* 1975; 55:604-613.

RILEY KF. *Mathematical Methods for the Physical Sciences.* pp301-5. Cambridge University Press, Cambridge (1974).

ROBBINS PA. The ventilatory response of the human respiratory system to sine waves of alveolar carbon dioxide and hypoxia. *J Physiol (London)* 1984; 350: 461-474.

SAHN SA, ZWILLICH CW, DICK N, MCCULLOUGH RE, LAKSHMINARYAN S, WEIL JV. Variability of ventilatory responses to hypoxia and hypercapnia. *J Appl Physiol* 1977; 43:1019-1025.

SARHAN NAS, LEANING MS, SAUNDERS KB, CARSON ER. Development of a complex model: breathing and its control in man. *Biomed Meas Infor Contr* 1987; 2:81-100.

SAUNDERS KB. Oscillation of arterial CO₂ tension in a respiratory model: some implications for the control of breathing. *J Theor Biol* 1980; 84:163-179.

SAUNDERS KB, BALI HN, CARSON ER. A breathing model of the respiratory system: The controlled system. *J Theor Biol* 1980; 84:135-161.

SEVERINGHAUS JW, BAINTON CR, CARCELEN A. Respiratory insensitivity to hypoxia in chronically hypoxic man. *Resp Physiol* 1966; 1:308-334.

STOCKLEY RA. The contribution of the reflex hypoxic drive to the hyperpnoea of exercise. *Resp Physiol* 1978; 35:79-87.

STOLL PJ. Respiratory system analysis based on sinusoidal variations of CO₂ in the inspired air. *J Appl Physiol* 1969; 27:389-399.

SUSKIND M, BRUCE RA, McDOWELL ME, YU PNG, LOVEJOY FW. Normal variations in end-tidal air and arterial blood carbon dioxide and oxygen tensions during moderate exercise. *J Appl Physiol* 1950; 3:282-294.

SWANSON GD, BELLVILLE JW. Hypoxic-hypercapnic interactions in human respiratory control. *J Appl Physiol* 1974; 36(4):480-487.

SWANSON GD, BELLVILLE JW. Step changes in end-tidal CO₂: Methods and implications. *J Appl Physiol* 1975; 39:377-385.

SWANSON GD, WHIPP BJ, KAUFMAN RD, AQLEH KA, WINTER B, BELLVILLE JW. Effect of hypercapnia on hypoxic ventilatory drive in normal and carotid body-resected man. *J Appl Physiol* 1978; 45:971-977.

TENNEY SM, ON LC. Ventilatory response of decorticate and decerebrate cats to hypoxia and CO₂. *Resp Physiol* 1977; 29:81-92.

TIBES U. Reflex inputs to the cardiovascular and respiratory centers from dynamically working canine muscles. *Circulation Respiration* 1977; 41: 332-431.

VON EULER US, LILJESTRAND G, ZOTTERMAN Y. The excitation mechanism of the chemoreceptors of the carotid body. *Skand Arch Physiol* 1939; 83:132-152.

WADE JG, LARSON CP Jr, HICKEY RF, EHRENFELD WK, SEVERINGHAUS JW. Effect of carotid endarterectomy on carotid chemoreceptor and baroreceptor function in man. *N Engl J Med* 1970; 282:823-829.

WARD SA, WHIPP BJ. Ventilatory control during exercise with increased external dead space. *J Appl Physiol* 1980; 48(2):225-239.

WASSERMAN K, WHIPP BJ, CASTAGNA J. Cardiodynamic hyperpnea: hyperpnea secondary to cardiac output increase. *J Appl Physiol* 1974; 36:457-464.

WEIL JV, BYRNE-QUINN E, SODAL IE, FRIESEN WO, UNDERHILL B, FILLEY GF, GROVER RF. Hypoxic ventilatory drive in normal man. *J Clin Invest* 1970; 49:1061-1072

WEIL JV, BYRNE-QUINN E, SODAL IE, KLINE JE, McCULLOUGH RE, FILLEY GF. Augmentation of chemosensitivity during mild exercise in normal man. *J Appl Physiol* 1972; 33:813-819.

WEIL JV, BYRNE-QUINN E, SODAL IE et al. Acquired attenuation of chemoreceptor function in chronically hypoxic man at high altitude. *J Clin Invest* 1971; 50:186-195.

WEISKOPF RB, GABEL RA. Depression of ventilation during hypoxia in man. *J Appl Physiol* 1975; 39:911-915.

WHIPP BJ, WASSERMAN K. Alveolar-arterial gas tension differences during graded exercise. *J Appl Physiol*. 1969; 27:361-365.

WHIPP BJ, WASSERMAN K. Carotid bodies and ventilatory control dynamics in man. *Fed Proc* 1980; 39:2628-2673.

WHIPP BJ, WIBERG DM. Eds. *Modeling and Control of Breathing*. Elsevier, New York (1983).

WHITE DP, DOUGLAS NJ, PICKETT CK, WEIL JV, ZWILLICH CW. Sexual influence on the control of breathing. *J Appl Physiol* 1983; 54:874-879.

WHITELAW WA, DERENNE J-P, MILIC-EMILI J. Occlusion pressure as a measure of respiratory centre output in conscious man. *Resp Physiol* 1975; 23:181-199.

WIBERG DM, BELLVILLE JW, BROVKO O, MAINE R, TAI TC. Modeling and parameter identification of the human respiratory system. *IEE Trans Auto Cont* 1979; AC-24(5):716-720.

YAMAMOTO WS. Mathematical analysis of the time course of alveolar carbon dioxide. *J Appl Physiol* 1960; 15:315-319.

YAMAMOTO WS, HORI T. Phasic air movement model of respiratory regulation and carbon dioxide balance. *Comput Biomed Res* 1971; 3:699-717.

ZWILLICH CW, SAHN SA, WEIL JV. Effects of hypermetabolism on ventilation and chemosensitivity. *J Clin Invest* 1977; 60:900-906.